



DANE

Para tomar decisiones

CONSUMER PRICE METHODOLOGY

COLLECTION OF DOCUMENTS – UPDATE 2009
Num. 62

NATIONAL ADMINISTRATIVE DEPARTMENT OF STATISTICS



JORGE BUSTAMANTE ROLDÁN
Director

CHRISTIAN JARAMILLO HERRERA
Deputy Director

MARIO CHAMIE
General Secretary

Technical Directors

NELCY ARAQUE GARCÍA
Regulation, Planning, Standardization and Normalization

EDUARDO EFRAÍN FREIRE DELGADO
Methodology and Statistical Production

LILIANA ACEVEDO ARENAS
Censuses and Demography

MIGUEL ÁNGEL CÁRDENAS CONTRERAS
Geostatistics

ANA VICTORIA VEGA ACEVEDO
Synthesis and National Accounts

CAROLINA GUTIÉRREZ HERNÁNDEZ
Diffusion, Marketing and Statistical Culture

Direction of Methodology and Statistical Production (DIMPE)

Eduardo Efraín Freire Delgado

Technical Coordination

María Ximena Caicedo Moreno (Coordinator), Clara Inés Gutiérrez Perdomo, María Teresa Rojas Linares, Jorge Andrés Tenorio Neira, Alexander González Coca, Vanesa Montoya Naranjo, Diana Marcela Sánchez Castro and Julieth Alejandra Solano Villa.

Technical team:

Statistical design; Irma Inés Parra Ramírez (Coordinator), Óscar Andrés Bautista Alfonso, Nancy Janneth Rodríguez Rodríguez. Systems design: Néstor Alonso Valbuena Gómez (Coordinator), Jakeline Suárez Cabezas. Logistics: Marisol Sabogal Hoyos (Coordinator).

Consulting Team:

Adolfo León Cobo Serna, Alejandro Reyes González, Carlos Alfonso Huertas Campos, José David Pulido Pescador, Héctor Manuel Zarate Solano (Banco de la República - Central bank of Colombia). Gabriel Armando Piraquive Galeano (National Planning Department) and Camilo Herrera Mora (Raddar).

International Consulting Team:

Louis Marc Ducharme (Assistant Chief Statistician, Business & Trade Statistics Field, Official Languages Champion) Statistics Canada (StatCan), and Jacob Ryten (former Assistant Chief Statistician) Statistics Canada (StatCan).

CONTENTS

	Page
PRESENTATION	10
INTRODUCTION	11
1. BACKGROUND	13
2. DESIGN	16
2.1. Conceptual framework	16
2.1.1. Objectives	16
2.1.2. Reference framework	16
2.2. Statistical design	20
2.2.1. Basic components	20
2.2.2. Design of processing and imputation methods.	39
2.2.3. Design of instruments.	40
3. PRELIMINARY ACTIVITIES	44
3.1. Awareness-raising process	44
3.2. Staff hiring	44
3.3. Collection of information	45
3.3.1. Operative organization Structure.	45
3.3.3. Mapping used.	47
3.3.4. Supervision and control of the operation of collection.	47
3.3.5. Classification and management of surveys or records.	49
3.4. Codification and analysis procedure	49
3.5. Capture and consolidation of data	49
3.5.1. Consolidation of files.	49
3.6. Processing and imputation	50
3.6.1. Coding method.	50
3.6.2. Adaptation of files.	50
3.6.3. Data editing.	51
3.6.4. Imputation criteria.	52
3.6.5. Data integration.	53

4. ANALYSIS	55
4.1. Statistical analysis	55
4.2. Analysis of quality of results	56
4.3. Quality indicators at territorial level	56
4.3.1. Response rate Index or coverage by sources (ITRF for its acronym in Spanish).	56
4.3.2. Non-imputation or local estimation Index (INIL for its acronym in Spanish).	56
4.3.3. Local Quality Index 1 (IDCL1 for its acronym in Spanish).	57
4.3.4. Local Quality Index 2 (IDCL2 for its acronym in Spanish).	57
4.4. Quality indicators at the central level	58
4.4.1. Response rate Index or coverage by records (ITRR for its acronym in Spanish).	58
4.4.2. Non-imputation Index or central estimate (INIC for its acronym in Spanish).	58
4.4.3. Reliability Index (ICFA).	59
5. DISSEMINATION	60
5.1. Data repository management	60
5.2. Dissemination products and instruments	60
5.2.1. Press release.	60
5.2.2. Press bulletin.	60
5.2.3. Presentation.	60
RELATED DOCUMENTATION	62
GLOSSARY	64
BIBLIOGRAPHY	67
ANNEXES	68
Annex A. Output Frames of the Consumer Price Index (CPI)	68

LIST OF TABLES

	Page
Table 1. Consumption and expenditure structure in two periods, 1 and 2	18
Table 2. Laspayres-type index	18
Table 3. Paasche-type index	18
Table 4. Collection period	21
Table 5. Structure by expenditure groups of the CPI-98 and CPI-08	26
Table 6. Structure by expenditure subgroups CPI-98, CPI-08	26
Table 7. Composition of the basket CPI-98, CPI-08	28
Table 8. Weights of the CPI-98 expenditure groups	29
Table 9. Weights of the CPI-08 expenditure groups	29
Table 10. Instructions for completing the form	40
Table 11. Category of establishments where more goods and services are purchased	46

LIST OF FIGURES

	Page
Figure 1. General, by income level for each city	30
Figure 2. General for the National Total	31
Figure 3. CPI capture form	51
Figure 4. Reports generated in the CPI application	51
Figure 5. CPI Environment of ANALYSIS	52
Figure 6. Reports, CPI environment of ANALYSIS	52

ACRONYMS

ACRONYM	DEFINITION IN ENGLISH	DEFINITION IN SPANISH
COICOP	Classification of Individual Consumption according to Purpose	Clasificación de Consumo Individual por Finalidades
COLI	Cost of Living Index	Índices de Costo de Vida
CPI	Consumer Price Index	Índice de Precios al Consumidor
DANE	National Administrative Department of Statistics	Departamento Administrativo Nacional de Estadística
DIMPE	Direction of Methodology and Statistical Production	Dirección de Metodología y Producción Estadística
DIVIPOLA	Political-Administrative Division	División Político Administrativa
DMC	Data Capture Device	Dispositivo Móvil de Captura
ENIG	National Income and Expenditure Survey	Encuesta Nacional de Ingresos y Gastos
EUROSTAT	Statistical Office of the European Communities	Statistical Office of the European Communities
FUR	Unique Collection Form	Formulario Único de Recolección
ICFA	Reliability Index	Índice de confiabilidad
IDCL	Local Quality Index	Índice de calidad Local
ILO	International Labor Organization	Organización Internacional del Trabajo
IMF	International Monetary Fund	Fondo Monetario Internacional
INIC	Non-imputation Index or central estimate	Índice de no imputación o estimación central
INIL	Non-imputation or local estimation index	Índice de no imputación o estimación local
ITRF	Response rate index or coverage by sources	Índice de tasa de respuesta o cobertura por fuentes
ITRR	Response rate Index or coverage by records	Índice de tasa de respuesta o cobertura por registros
NPISH	Non-for-Profit Institutions that Serve the Homes	Instituciones Sin Fines de Lucro que Sirven a los Hogares
NSS	National Statistical System	Sistema Nacional de Estadística
OECD	Organization for Economic Cooperation and Development	Organización para la Cooperación y el Desarrollo Económico

PE	Waiting Period	Período de Espera
SIPSA	Price Information System for the Agricultural Sector	Sistema de Información de Precios del Sector Agropecuario
UN	United Nations	Organización de las Naciones unidas
VAT	Value Added Tax	Impuesto al Valor Agregado
WB	World Bank	Banco Mundial

PRESENTATION

The National Administrative Department of Statistics (DANE), in its role as coordinating entity of the National Statistical System (NSS), and within the framework of the “Statistical Planning and Harmonization” project, works towards the strengthening and consolidation of the NSS. This is carried out through several processes, such as: the production of strategic statistics; the generation, adaptation, adoption and dissemination of standards; the consolidation and harmonization of statistical information, and the connection of instruments, stakeholders, initiatives and products. These actions are carried out in order to improve the quality of strategic statistical information, and its availability, timeliness and accessibility, so as to respond to the users’ demand.

In this context and being aware of the need and obligation to provide better products to its users, DANE developed standard guidelines for the presentation of methodologies, to contribute to the visualization and understanding of the statistical process. By means of this instrument the entity prepared methodological documents on its operations and statistical research that are made available to specialized users and the general public. The documents present the main technical characteristics involved in the processes and sub-processes of each research in a standard and comprehensive manner, thus allowing further analysis, control, replicability and assessment.

These series of documents promote transparency, trust and credibility of the institution’s technical expertise, for a better understanding and use of the statistical information that is produced in accordance with the principles of coherence, comparability, integrality and quality of statistics.

INTRODUCTION

The Consumer Price Index (CPI) is a statistical research that measures the average percentage change in prices of a set of final goods and services that consumers demand.

CPI information enables the analysis of economic situations and is used for decision-making by the Government and private entities. The index is primarily used as an adjustment factor in the determination of wages, financial statements, and in the solution of labor or fiscal claims.

The CPI is used to calculate the currency's loss of purchasing power, to obtain equilibrium in national accounts items, and as an analysis factor of the behavior of the economy.

The first works on the Consumer Price Index were completed by the Banco de la República (Central bank of Colombia) in 1923. In the same year, the Office of the Comptroller General of Colombia was created; this entity assumed the function of structuring official statistics and in 1937 developed the first works on price indices, on the basis of expenditure surveys of the working class in the city of Bogotá.

Subsequently, surveys were carried out in the cities of: Medellín, (with May of 1983 base), Bucaramanga (with September 1945 base), Barranquilla (with October 1946 base), and Manizales (with October 1947 base). In 1954 DANE takes over the responsibility of the production of the index, and since then, the methodological, conceptual, and construction aspects for the index were established. These aspects could be considered as being more standardized, reflecting an institutional effort of permanent renovation to offer a more realistic approach of the statistical exercise to the price effect that affects the Colombians' final consumption expenditure.

According to Decree 3167 dated December 26 of 1968, the production of the CPI was explicitly established within the DANE's functions, as follows: "To establish price indices, at producer, distributor, and consumer levels, on the main goods and services; conduct surveys and periodically publish the summary of the results obtained" (Chapter I, Article 2, paragraph J).

DANE reviews the CPI regularly aiming at the incorporation of methodological and performance variables that make the statistical production of the index more complete and in accordance with the new developments contributed by the more advanced statistical systems on this matter. According to international recommendations and standards, the CPI must be updated at least every 10 years. In compliance with these standards, DANE made the most recent update for Colombia in the year 2008.

For the CPI updating process, the standards, recommendations and norms established by the following entities were used as reference (regulatory framework of statistical entities): Governing bodies, United Nations (UN) and the International Labor Organization (ILO), International agencies or organizations, the International Monetary Fund (IMF), and general practices of prestigious and traditional international statistics institutes from Canada, the United States, France, as well as national and international experts on the subject (Statistics Canada).

As basic documents, and with reference to the best practices, for this update the Consumer Price Index Manual “Practice and Theory”, by the International Labor Organization (ILO), the International Monetary Fund (IMF), the World Bank (WB), the Organization for Economic Cooperation and Development (OECD), the Data Quality Assessment Framework from the International Monetary Fund (IMF), the Norm for the Calculation of a Harmonized Index of Consumer Prices for the Member Countries of the Andean Community, and the Manual of Good Practices of EUROSTAT, were used.

As a result of the mentioned reviewing process, DANE delivered a new version of the index to the country, which comprises the updating of the market basket for families and weighted calculations, as well as the increase in the geographical coverage of the indicator, based on the most recent National Income and Expenditure Survey, conducted in the years 2006-2007.

This document pretends to be a practical guide that compiles the main information contained in the methodologies (thematic, statistical, sampling, and of indicators) and in other documents that are part of the Quality Management System and of the Process Documentary System of the Consumer Price Index (CPI) research.

1

BACKGROUND

Some of the initial works on CPI date to 1917. Office of the Comptroller General of Colombia did some works on retail prices after 1917. As from the year 1950, DANE is strengthened in the production of price indicators, thanks to the recommendations of the economic missions on the consolidation of the planning system. Since 1954 five revisions to the production of the indicator have been included, with updating of weights, extension of geographical coverage and selection of the market basket for the monitoring of prices, namely:

CPI-20

Basic survey: household Income and Expenditure surveys carried out in August 1953 in the city of Bogotá, and in October of 1953 in the city of Medellín, Cali, Barranquilla, Bucaramanga, Manizales, and Pasto.

Base period and validity of the review: base: July 1954- June 1955 = 100, from July 1954 to December 1978.

Coverage of the index

a. **Geographical:** makes reference to the population of seven cities, namely: Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales and Pasto. The criteria used for the selection of these cities were: level of income, population density, weather, and the location within the geographical circuit of the country.

b. **Socioeconomic:** households in the surveyed cities were stratified in two main occupational classifications: those whose head of household was an employee and those whose head of household was a worker.

The following aspects were considered for the socioeconomic classification: The persons' occupational category, the occupational position of family members determined by the presence of physical or intellectual work, the training required for the performance of their duties, functions, and the standard of living measured by dwelling conditions.

Independent workers, non-family groups, and households constituted by adults only were excluded from this classification process.

CPI-40

Basic survey: the consumption structure used to establish the CPI weights is observed in the Income and Expenditure Survey carried out in 1970, in the country's seven major cities, and revised to take into account the evolution of prices as from this date until the base period.

Base period and validity: base December 1978 = 100, from January 1979 to December 1988.

Index coverage

- a. *Geographical*: the information is representative of the population of seven cities: Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales and Pasto.
- b. *Socioeconomic*: single-person households, collective households, and high-income households were excluded from the reference population. The Income and Expenditure Survey of 1970 revealed that the structures of consumption depended, to a greater extent, on the household income level, more than on the occupational categories in which households were classified, consequently, the previous worker and employee classification was substituted by the low-income and medium-income households' classification, excluding high-income households.

CPI-60

Basic survey: the Income and Expenditure Survey carried out between March of 1984 and February of 1985 constitutes the basis for the establishment of the CPI-60. In this survey the weekly expenditures were reported daily by housewives, by means of the “*diary method*” that was used for the first time in investigations carried out by DANE. The survey covered 52 collection weeks and was carried out in 15 cities.

Index coverage

- a. *Geographical*: the information is collected in the following 13 cities: Bogotá, Medellín, Cali, Barranquilla, Bucaramanga, Manizales, Pasto, Pereira, Cúcuta, Montería, Neiva, Cartagena and Villavicencio, with their respective metropolitan areas or zones of influence. This geographical coverage is representative of the urban population and strengthens the representativeness of the urban national index.
- b. *Socioeconomic*: the reference population is defined by the set of private households located in the areas of the thirteen cities. Single-person, collective, and high-income households were excluded. Taking into account that the households' purchasing power constitutes a determinant variable in the consumption structure, the CPI-60 continues to use the two groups differentiated by low-income and medium-income households.

Base period and validity: base December 1988=100, from January 1989 to December 1998.

CPI-98

Basic survey: the Income and Expenditure Survey for the CPI-98 was carried out between March 1994 and February 1995. For the construction of the market basket the values were adjusted according to the evolution of prices of the goods and services included, during the period between the survey collection and December 1998.

Base period and validity: the base of the index is constituted by the average prices of the month of December 1998, meaning that the index will be December 1998=100.

Index coverage

- a. *Geographical*: this refers to the population of 13 department capitals, and includes some metropolitan areas. The cities with local indices are: Bogotá; Medellín and Bello; Envigado and Itagüí; Cali and Yumbo; Barranquilla and Soledad; Bucaramanga with Floridablanca, and Girón, Manizales and Villa María; Pasto; Pereira and Dosquebradas; Cúcuta; Los Patios; El Zulia and Villa del Rosario; Cartagena; Montería; Neiva and Villavicencio.

b. *Socioeconomic*: reference households are defined by the set of private households located in the urban area of thirteen cities. Single-person households and collective households were excluded.

The CPI-98 defined three groups differentiated by income, that is, low, medium, and high-income households.

The CPI-98 is the index revision that included the most changes; this is why it should be considered as the most significant variation of the last years. From this revision the following stand out: The change in the index's production and dissemination systems (including a fixed-level and another flexible level that enables a faster update of the basket for price-tracking purposes), the calculation procedure (using the geometric reason) and the use of implicit models for quality adjustment.

CPI-08

The CPI-08 is the most recent revision of the index; it comprises an update of the market basket and calculation weights, as well as the increase in geographical coverage of the indicator, based on the most recent 2006-2007 income and expenditure survey. In addition, it includes a rationalization in the number of basic expenditure and in the articles for price-tracking purposes.

The technical details of this revision are the fundamental body of the methodology that is presented below.

2 DESIGN

2.1. CONCEPTUAL FRAMEWORK

2.1.1. Objectives

General objective: to collect and present, from a base month, the monthly average change in prices of a basket of goods and services representative of household consumption in the country, with the purpose of strategically guiding the decision-making process and analysis of economic situations, both of the Government and of private entities.

Specific Objectives

To provide statistical data that enable to perform the updating processes of monetary flows and indexation.

To provide data for the deflation of nominal values, thereby determining the evolution in real terms of an economic item

To provide data to enable economic and social situational analysis.

2.1.2. Reference framework

Conceptual base. As part of the design of an indicator that reflects the consumer prices' evolution, it must be decided between the construction of a constant utility index, theoretically known as Cost of Living Index (COLI), or a price indicator with a fixed market basket and utility. In practice, the basket and fixed-income indices are considered a good approximation to the Cost of Living Indices (COLI).

Cost of living index: in developing a Cost of Living Index (COLI) two different types of utility functions should be considered; first, each individual consumer is set as an agent, while the second one refers to the aggregate of consumers. In practice, it is difficult to build these two utility functions as from the traditional statistical tools (surveys).

In a COLI the basket should consider all the goods and services that are part of the households' expenditure structure, without exception. This implies that it is not possible to apply selection criteria to the basket from the goods' and services' characteristics, the forms of acquisition, or payment mechanisms.

Within the framework of a COLI the demand function depends on prices, quantities, and utilities:

$$COLI = f(\text{Pr}, Q, U_i)$$

Where:

Pr: Prices

Q: Quantities

U_i: Utility

Bearing in mind the above, the development of a COLI is aimed at evidencing the minimum expenditure change that is necessary to purchase a basket of goods and services, keeping the level of utility or wellness obtained by the consumer, all this comparing and analyzing two periods of time.

Consumer Price Index: in practice, the development of this type of indicator emerges as an alternative, which approximates to the COLI. For this, a consumer-preference structure is established, which remains fixed. Thus, the problem of constructing an indicator of consumer prices is reduced to the measurement of quantities and prices.

The preferences' structure used corresponds to that observed in the households' consumption and expenditure structure, in a given period of time (one can choose between a base period, or a current period).

In the case of consumer price indices, those goods and services for which it is possible to identify their acquisition prices, brands, quality, and price monitoring units, that is, end-consumer goods are included in the index. This enables the definition of the scope of application of the research on prices and distinguishes it from the spectrum considered by the Cost of Living Index.

The success of a fixed-basket index is achieved by effectively representing expenditure patterns and household consumption from the goods and services selected to be part of the basket, for price-tracking purposes. In this sense, aspects related to the population under study and geographical coverage of the exercise are also considered important.

Type of Composite Index: the alternatives in terms of calculation formulas are among the use of three types of composite indices, namely: A fixed-weight index number for the base period of the exercise, another with variable weights related to the current period and, finally, a combination of these two, or superlative. This index composition enables minimization of implicit biases in the statistical exercise on consumer prices with fixed baskets, and meeting the objective of the indicator, which is reporting price changes that affect the consumer.

Fixed-weight indices: if two time periods are considered, a Laspeyres-type index (theoretical designation of a fixed-weight index) aims to answer the question: ¿How much does the purchase price of a basket of goods and services increase or decrease, if purchase quantities from the first period are kept fixed? These fixed quantities are related to the expenditure and consumption pattern of the exercise's base period and, hence, the preference structure evidenced by households. Changes that may occur in the purchasing value are assumed as effective variations in prices.

In the theory of indices, it is considered that Laspeyres-type indices may present an upward substitution bias in relation to the constant utility indices, because they do not consider the fact that households can perform substitution processes from changes in relative prices, or changes in income level.

Variable-weight indices: with a variable-weight index or Paasche-type index, for two periods of time, we attempt to answer the following question: ¿How much does the purchase price of a basket of goods and services increase or decrease if the quantity bought today could be purchased maintaining the prices from the initial period?

In variable-weight indices the weights' structure and, consequently, the household preferences structure correspond to the current period. In this sense it is possible to associate changes in the value of the basket with effective price variations.

In the indices theory, contrary to what happens with fixed-weight indices, it is considered that Paasche-type indices may present a substitution bias downwards, since the reference point for calculating the variation is the consumption structure of the current period, and given that this structure reflects the price conditions of the current period and not the consumption structure of a former period, the price variations observed may be lower.

Superlative index or Fisher ideal: Fisher indices emerge as an alternate solution to the substitution biases that may occur in the Laspeyres and Paasche price indices. In this sense they are also approximations to the Cost of Living Index. Thus, with a Fisher ideal index one seeks to incorporate the price changes that are adequately reflected in a Laspeyres-type index and changes in the preferences structure or consumption patterns incorporated into a Paasche-type index. Given that with a Fisher index the above-mentioned upwards and downwards substitution biases are resolved, they are also known as superlative indices or Fisher ideals.

In practice, for the construction of the superlative index the information requirements turn out to be greater than in the individual case of the other indices; it is necessary to build upon them, both Laspeyres-type and Paasche-type price indices.

Table 1. Consumption and expenditure structure in two periods, 1 and 2

Item	Q1	P1	V1	Q2	P2	V2
Beef	1,0	1,0	1,0	0,8	1,6	1,28
Chicken	1,0	1,0	1,0	2,0	0,8	1,60
Total			2,0			2,88

Source: Exercise cited in the Boskin report, "Toward A More Accurate Measure Of The Cost Of Living" issued on December 4, 1996 by the Advisory Commission To Study The Consumer Price Index. United States Congress

Table 2: Laspeyres-type index

Item	Q1	P1	V1	Q1	P2	V est.	Variation
Meat	1,0	1,0	1,0	1,0	1,6	1,6	
Chicken	1,0	1,0	1,0	1,0	0,8	1,6	
Total			2,0			2,4	1,2

Source: Exercise cited in the Boskin report "Toward a More Accurate Measure of the Cost Of Living" issued on December 4, 1996 by the Advisory Commission to Study the Consumer Price Index. United States Congress

Table 3: Paasche-type index

Item	Q2	P1	V1	Q2	P2	V est.	Variation
Beef	0,8	1,0	0,8	0,8	1,6	1,3	
Chicken	2,0	1,0	2,0	2,0	0,8	1,6	
Total			2,8			2,9	1,0

Source: Exercise cited in the Boskin report "Toward a More Accurate Measure of the Cost Of Living" issued on December 4, 1996 by the Advisory Commission to Study the Consumer Price Index. United States Congress.

Superlative Index or Fisher Ideal:

$$ISF = ITL * ITP = 1,20 * 1,03 = 1,11$$

Where:

ITL= Laspeyres-type index.

ITP= Paasche-type index.

International Benchmarks: at international level, most indexes are calculated taking into account the reference framework of the Consumer Price Index Manual provided by the ILO and the IMF, which refers to the calculation methodology for an index, uses of indices, concepts and scope, expenditure weights and sources, sampling, and collection of prices (International Monetary Fund, 2006).

Another reference framework is the IMF's Statistics Department, in charge of the quality assessment of the CPI, which is related to the following components: Preconditions for quality, integrity, methodological rigor, accuracy and reliability, and usefulness of statistics for the user.

Since 1995 several countries, Colombia among them, have taken into consideration the recommendations of the U.S. Boskin report for the CPI ("Toward a More Accurate Measure of the Cost Of Living" Dec. 1996), given that it is a comprehensive evaluation of the design or operation biases which may be incurred in the design and operation of a CPI (Boskin, 1996).

It is important to note that DANE has followed the recommendations made by these international benchmarks.

Thematic content

a. *Classification variable:* groups, subgroups, classes and basic expenditures, income levels and geographic area.

b. *Analysis variables:* price and variation in end-consumer prices of goods and services.

c. *Calculated variable:* geometric average of price variations and weighted average of index numbers.

d. *Indicators:*

- Simple relatives by source,
- Geometric average of simple indices by source for each item or variety,
- Weighted average of geometric averages of simple indices per item or variety for each basic expenditure,
- Local basic expenditure index,
- National basic expenditure index per income levels,
- Total basic expenditure index,
- Expenditure class index,
- Expenditure subgroup index,

- Expenditure group index,
- Total index as an average of expenditure groups,
- Variation.

2.2. STATISTICAL DESIGN

2.2.1. Basic components

Type of statistical operation

Survey by sampling:

Universe: consists of all the establishments where the consumer turns to, in order to purchase goods and services for consumption. This includes retail establishments, real estate or informal services of rental of dwellings, schools, public utilities companies, health providers, etc.

For the specific case of education, the universe is conformed by official and non-official schools that provide preschool, elementary, middle, and high-school education, as well as by formal and non-formal higher education establishments.

To capture price variations in rented and owner-occupied dwellings, all of the existing rented units in the 24 cities of the investigation are considered.

Target population: all the establishments where the consumer purchases goods and services for consumption. This includes retail establishments, rented dwellings, schools, utilities' companies, etc.

Coverage and geographical disaggregation: the CPI research covers the population of 24 department capitals, including some metropolitan areas, which account for the largest number of habitants in the country and have the highest population and economic dynamics with respect to other cities.

These cities are: Bogotá; Medellín, Bello, Envigado and Itagüí; Cali and Yumbo; Barranquilla and Soledad; Bucaramanga, Floridablanca, Piedecuesta and Girón; Manizales and Villa María; Pasto; Pereira and Dosquebradas; Cúcuta, Los Patios, El Zulia and Villa del Rosario; Montería; Neiva; Cartagena; Villavicencio; Riohacha; Armenia; Quibdó; Sincelejo; Valledupar; Popayán; Ibagué; San Andrés; Santa Marta; Tunja, and Florencia.

Observation unit: the observation units are the establishments where the consumer purchases goods and services for consumption.

Unit of analysis: the unit of analysis refers to the prices of the goods and services purchased by consumers.

Sampling unit: sampling units are the establishments where the consumer purchases goods and services for consumption.

Nomenclatures and classifications: the current classification of goods and services that are part of the CPI research is based on the structure used by National Accounts and in the classification used in the CPI-60 and the CPI-98. It is worth mentioning that the latter classification keeps the guidelines of the United Nations Classification of Individual Consumption according to Purpose (COICOP)¹.

¹ The COICOP classification published by the Statistic Division of United Nations groups the households' consumption expenditures in three institutional sectors: households, Non-for-Profit Institutions that Serve the Homes (NPISH) and the Government.

Each of the cities under study is codified in accordance with the Political-Administrative Division (DIVIPOLA for its acronym in Spanish). This classification structure enables maintenance of the historical continuity of the research information.

The classification adopted by the CPI-08 includes nine groups of expenditure, given that the former transportation and communications group was split in two: Transportation itself and the Communications Group. The nine groups are:

1. Food products
2. Dwelling
3. Apparel
4. Health
5. Education
6. Culture, entertainment and recreation
7. Transportation
8. Communications
9. Other expenses

Reference period: the reference period for the research is *one month*.

Collection period: this is related to the frequency with which prices are modified; thus, a series of periods for collection have been defined, namely: Monthly, bimonthly, quarterly, every four months, annually, and others with open periodicity. Those items whose prices can change at any time are included in the open periodicity, and should be followed-up throughout the month (Table 4).

The sample should be evenly distributed during the days of each collection period, taking into account its source group, source class, and item.

National conformation prices are those that are taken exclusively in the city of Bogotá, since companies establish these prices for all the cities in a centralized manner. Within this category we find the national newspaper, long-distance national telephone service, and wireless telephony, among others.

Table 4. Collection period

Periodicity	Good or service
Monthly	Food
Bimonthly	Household goods
	Toiletries
	Medicine
	Personal items
	Liquors
Quarterly	Wardrobe
	Appliances

Table 4. Collection period

		Conclusion
Periodicity	Good or service	
Every four months	Furniture	
	Health services	
	Remainder	
	Rentals	
Semiannual	Higher education	
Annual	Schools	
Open	Public services	
	City transport	
	Municipal transport	
	Gas	
	Newspapers	
	Lottery	
	Soccer	
	Airline tickets	

Source: DANE

CPI basket.

Basic information to construct the CPI basket: The basket for price-tracking purposes and the structure of weights are obtained from a specialized survey called the National Income and Expenditure Survey (ENIG for its acronym in Spanish), that is regularly conducted by DANE.

This Income and Expenditure Survey applied by DANE in several periods (1970, 1984-1985, 1994-1995, and the most recent one, 2006 - 2007), enables to be certain as to the expenditure and consumption patterns of many families in the country, given that it inquires about the daily expenditures of each household member and the periodicity of these expenditures (weekly, monthly, quarterly, semiannually, and annually). From this information a general consumption pattern for households in the country can be constructed.

The ENIG² carried out in the years 2006 and 2007 captured information from rural areas, and introduced the *purchase place*, variable, which is of great use in the tracking of prices carried out by the CPI given that it provides information on households' consumption, habits and about the establishments where they use to purchase their goods and services. This inclusion takes into account new forms of purchasing, related to technological and marketing innovations, such as online stores, catalog sales, and specialized trade fairs.

In this survey the information that enabled to detect changes in the household's purchasing habits, related to the purchase periodicity of certain items, through the inclusion of the *Frequency of purchase* variable in forms two (2) *urban and rural*, and form three (3) *notebook of less frequent expenses*, was also captured; in addition, the chapter *characteristics of household's food expenditure* was included for the urban and rural areas.

Likewise, the CPI also uses information from the ENIG to know the relative importance (weight) of the different items in the monthly household expenditure, to define the goods to be included in the fixed and flexible structure of the index, and determine to which goods and services is price-tracking going to be applied to.

² The general objective of the ENIG is to obtain information on the amount or distribution of expenditures of Colombian households, as well as the amount and sources of income. The information on households' income and expenditures is gathered by means of a direct survey on forms two (2) rural and form three (3) less frequent expenses. Moreover the "diary method" is used for forms two (2) urban and four (4) urban.

The information thus gathered enables to affirm that no expenditure or consumption pattern of a particular household in the country is the same as another. Although the list of goods and services demanded by consumers is largely the same for all households, each consumer buys a different combination thereof, in quantity and frequency of time.

However, as stated earlier, due to the aggregation of individual patterns a generalized pattern of domestic consumption is obtained, which enables the selection of the basket for price-tracking and also to obtain the structure of expenditure weights (price by quantity). This pattern of generalized consumption results in the production of consolidated information on the behavior of prices, based on a wide range of goods and services that make up the CPI.

Criteria for the selection of the Basket for price-tracking purposes: The goods and services included in the basket for price-tracking of the CPI-08 were selected by applying general and specific criteria on the expenditure information reported by the country's households.

Among the general criteria the following can be mentioned:

- Items must have characteristics that facilitate price-tracking.
- Items must offer sufficient guarantee of permanence in the market.

On their hand, the particular criteria applied on the ENIG, sorted by cities and income levels, are described below:

Expenditure share

Different forms of validation are presented for the application or usage of the criterion. In some countries, the behavior or evolution of the participation among different household income and expenditure surveys is analyzed, in order to determine the inclusion or exclusion of an item in the basket for price-tracking purposes,³ depending on the profit or loss in percentage of the expenditure and on its growth expectations. In other countries, an absolute value in the participation is determined, from which the items that are part of the basket are classified for price-tracking purposes. It is noteworthy that such participation is defined arbitrarily and is applied with different parameters in each statistical system that is considered.

In the second case, some statistical agencies take as a parameter the participation in total consumption expenditure, with differential levels depending on the price variability for large aggregates, that is, participation for food products and participation for items other than food products⁴. According to international standards, a general principle for the construction of the basket is the participation in spending of less aggregated sub-groups of categories, like the class of expenditure or the basic expenditure. This enables the sample to be more representative of the price effect at those levels (International Monetary Fund, 2006).

In the case of Colombia for the CPI-08, the reference on the expenditure share as a selection parameter is determined by a 5 % of the expenditure within the subclass and a 0,01 % compared to the total expenditure.

³ CPI of Canada applies this criterion.

⁴ In the Canadian case, the determined participations are 0,1 % of the total expenditure for food products and 0,2 % for non-food products. In Spain the participation is 0,03 % of the total expenditure.

Frequency of demand

The frequency of the demand is used as a selection criterion, especially to determine the *forced inclusion* of an item. Most statistical systems do not make explicit the use or definition of values for the frequency of demand. In the case of Colombia's CPI, since the CPI-40 methodology⁵, the frequency of demand has been taken into consideration and a population percentage has been used as a classification parameter. Indeed, if an item has a *frequency of demand* that exceeds 30% of consumer households⁶, then it must be included in the basket, regardless of its relative weight in the expenditure.

Analysis of the evolution of the share in the consumption expenditure

Specifically, the evolution occurring between ENIG 1994-1995 and ENIG 2006-2007 has enabled the selection of goods and services, whose growth dynamics in the share of consumption expenditure is important, even if those goods and services did not classify according to the criteria of expenditure share.

Expected growth in the demand for an item or product

This particular criterion enables the inclusion of products that appear in the income and expenditure survey for the first time, and which have a low expenditure share and frequency of demand. Nevertheless, these products present interesting growth expectations in the short and medium terms, therefore, not considering them would eventually and rapidly cause outdated of the basket for price-tracking purposes.

Structure of the CPI basket: before presenting the basket structure it is important to bear in mind that the articles' classification system used by the ENIG follows the *Classification of Individual Consumption According to Purpose (COICOP)* scheme, whose main classification criterion is the article's origin. For the basket selection, weights calculation, and CPI contrast and evolution exercises it was necessary to relate the COICOP system to the CPI system.

In the CPI system the selected goods and services are organized in accordance to a classification system or construction structure system, enabling each good or service to occupy a unique place in it. Additionally, said classification system makes it possible to obtain the national total as a result of the aggregation of categories, or the simpler levels that are presented below:

- Expenditure groups – aggregation of expenditure subgroups.
- Expenditure subgroups – aggregation of expenditure classes.
- Expenditure classes – aggregation of basic expenditure.
- Basic expenditures – aggregation of items or varieties.

Criteria for the construction of the general structure: in order to define the general structure of the index, as from the revision of the CPI-98 and also for the CPI-08, the following international recommendations have been taken into consideration:

General criteria and conditions

- The classification of goods and services to be used corresponds to the *classification of the households' consumption* used in *National Accounts*, specifically the final usages presented therein.

⁵ Used by DANE from 1974 until 1988.

⁶ According to households that responded to the income and expenditure survey.

- The selected structure assures the continuity that the index has been producing in the statistical information, until a reasonable level of usefulness for the project's main and general users is attained.
- The selected structure must allow a certain degree of international comparability of the statistical information produced by the index, at least with those statistical systems where the countries' social and economic conditions make the comparison feasible.
- The selected structure should enable minimization of bias emergence in the measurement of the price effect associated to the households' final consumption expenditure, biases that are inherent to the work of consumer price indices as from fixed baskets.

Particular criteria

- Use functions have been identified from the consumers' perspective, for the conformation of aggregates.
- *Expenditure substitution or complementarity processes* have been identified for the construction of aggregates.

Components of the new CPI structure: since the revision of the CPI that was carried out in 1998, DANE made progress in a methodological variant about the Laspeyres-type indices, consisting of the identification of two general levels in the structure, with differences in their characteristics and updating method. These levels are: Firstly, a *fixed* component of the structure; and secondly a *flexible* component, which implies significant differences with regard to the previous CPI structures.

Fixed level

Characteristics and updating method:

- The CPI's *fixed* level is the part of the indicator structure that continues to be used in the traditional framework of the Laspeyres-type indices, i.e., it has a *fixed expenditure weighting* derived from the latest ENIG.
- The CPI's fixed level is the part of the indicator structure that will be subject to dissemination and publication.

The CPI's fixed level is the part of the indicator structure that can only be updated from the application and processing of a new ENIG. This means that the level remains unchanged during the life of the CPI revision.

Fixed-level components

The CPI's fixed level comprises four categories. From the highest to the lowest level of aggregation, namely: Expenditure group, expenditure subgroup, expenditure class, and basic expenditure.

Expenditure group: This is the most general level of the structure; the updated version of the index keeps the concepts and nature of the expenditure groups of the CPI-98. The only difference is the emergence of a new expenditure group, which is the result of the division of the "Transportation and communications" group into two groups; one is "Transportation" and the other is "Communications"⁷.

⁷ The division of the "Transportation and Communications" group into "Transportation" and "Communications" was due to the fact that differences in the behavior of prices between the two group components were evidenced in recent years, as well as in the expenditure share registered by the communications component in the 2006-2007 ENIG.

Table 5. Structure by expenditure groups of the CPI-98 and CPI-08

ICPI 1998	CPI 2008
1. Food products	1. Food products
2. Dwelling	2. Dwelling
3. Apparel	3. Apparel
4. Health	4. Health
5. Education	5. Education
6. Entertainment, culture and recreation	6. Entertainment, culture and recreation
7. Transportation and communications	7. Transportation
8. Other expenses	8. Communications
	9. Other expenses

Source: DANE

- Expenditure subgroup: in the following level of aggregation there is a conceptual equivalence between the CPI-98 and the CPI-08.

The division of the “Transportation and communications” group produces a code change in the corresponding expenditure subgroups.

Table 6. Structure by expenditure subgroups CPI-98, CPI-08

Code	CPI-1998	Code	CPI-2008
11	Cereals and bakery products	11	Cereals and bakery products
12	Tubers and plantains	12	Tubers and plantains
13	Vegetables and legumes	13	Vegetables and legumes
14	Fruits	14	Fruits
15	Meat and meat products	15	Meat and meat products
16	Fish and other sea products	16	Fish and other sea products
17	Milk, eggs and fats	17	Dairy products, fats and eggs
18	Other foods	18	Other foods
19	Meals outside the home	19	Meals outside the home
21	Dwelling occupancy expenses	21	Dwelling occupancy expenses
22	Fuels and utilities	22	Fuels and utilities
23	Household furnishings	23	Household furnishings
24	Household appliances	24	Household appliances and repair of household appliances
25	Household utensils	25	Household utensils
26	Household linens	26	Household linens
27	Items for household cleaning	27	Items for household cleaning
31	Apparel	31	Apparel
32	Footwear	32	Footwear
33	Apparel and footwear services	33	Apparel and footwear services
42	Goods and health items	42	Goods and health items
41	Health Professional Services	41	Health services
43	Private insurance costs	43	Private and social insurance costs
51	Instruction and teaching	51	Instruction and Teaching
52	School supplies	52	School supplies

Table 6. Structure by expenditure subgroups CPI-98, CPI-08

				Conclusion
Code	CPI-1998	Code	CPI-2008	
61	Cultural and other related articles	61	Cultural items and others related	
62	Apparatuses for fun and leisure activities	62	Apparatuses for fun and leisure activities	
63	Services, hobbies, entertainment and leisure activities	63	Services for fun and leisure activities	
71	Personal transportation	71	Personal transportation	
72	Public transportation	72	Public transportations	
73	Communications	81	Communications	
81	Alcoholic beverages and cigarettes	91	Alcoholic beverages and cigarettes	
82	Toiletries and personal care	92	Toiletries and personal care	
83	Jewelry and other personal items	93	Jewelry and other personal items	
84	Other goods and services	94	Other goods and services	

Source: DANE

- **Expenditure class:** the identification of expenditure classes aims to capture the combined price effect of very similar categories, or that might work as a substitute, associated with the use functions.
- **Basic expenditures:** this is the fundamental level of the structure, and captures a very pure price effect. It also corresponds to the lowest point which has a fixed weighting. As it will be seen hereinafter, in some cases, this can be assimilated to an item level.

Flexible level

As from the CPI-98 version, the CPI has a flexible structure, ergo, subject to be modified based on a specialized economic and statistical analysis, which enables the detection of changes in the patterns that are used for price-tracking purposes.

The highlighted flexibility enables the quoting of prices on a wider range of goods and services, thus capturing the heterogeneity of demand among the country's individuals and regions. It also enables a faster updating of the pattern for price-tracking purposes.

Another advantage of working with a flexible structure is the use of the *geometric average* to calculate simple indices, making explicit the processes of substitution of items, or between varieties of items on the part of the consumer, in addition to having the mathematical consistency on paired samples, whereby the percentage variation in price averages, or the average percentage variations, offer the same result.

Having a flexible level enables capturing or incorporating the price effect that the emergence of a new item entails, even without having, at this level, a weighted expenditure (built from ENIG). To do this, we proceed in the framework of an expenditure reweighting among items belonging to the basic expenditure, where the new good or service could be located.

Flexible level components

The flexible level in the new CPI may be composed of items or varieties of items, depending on their relative importance within the expenditure, or within the consumption pattern of families in the country, from where the basic expenditures may have the following characteristics:

- **Basic expenditure defined as an item in general:** within basic expenditures, this situation comprises the identification of items, which expenditure importance, frequency of demand and generalized importance in all the cities that are part of the geographical framework⁸ of this exercise, make them worthy of this category, among them: rice, potatoes, bread, milk, cheese, medical consultations, etc.

For these, the price variation calculation focuses on capturing varieties of the same product that might be presenting a substitution in the matter of expenditure, or a complement of it.

- **Basic expenditure defined as an aggregation of goods and services:** this situation comprises the identification of some expenditure items that lack the sufficient importance to be individually constituted as basic expenditure; however, price-tracking of these goods and services is of interest in the short or medium terms. Because of this, basic expenditures are formed from grouping goods and services that are homogeneous in their characteristics, or price behavior.

For the above-mentioned basic expenditure, the price variation is obtained from the average of variations, first by item and subsequently these variations are aggregated up to the basic expenditure level.

Under this situation, the way of capturing regional consumption patterns may be lumped together, placing the basic expenditure as “Others”, and below this level to handle the differential of such consumption patterns.

Thus, the presence of items in the national structure of weights can be avoided, since the expenditure importance of these items is minimal and more weights appear due to regional considerations. In other words, said weights are important in the consumption of some cities from the geographical framework, but do not enable a perfect consistency in terms of weights to obtain the indicator in the aggregate level, or for the national total.

A similar situation arises with respect to the difference in consumption patterns by income levels, for which the differences are not so sharp. That is, in terms of the structure, it is not that a number of completely different goods and services are consumed, but that the proportions of expenditure differ. And although the price levels of the same goods and services are different, the variations and their tendencies have the same behavior. Against this for the tracking of prices it is possible to have different items by income levels (low, medium, high) at the basic expenditure level “Others”⁹.

The composition of the basket, designed based on the application of the above criteria, is presented in Table 7.

Table 7. Composition of the basket CPI-98, CPI-08

Structure level	CPI 08	CPI 98
Groups	9	8
Subgroups	34	34
Classes	88	79
Basic expenditures	181	179
Items or varieties	423	408

Source: DANE

⁸ Becoming almost a national consumption pattern.

⁹ It should be noted that the expenditure importance is higher in some articles than in others.

Basket weights

To obtain the relative weight of each good and service within the CPI basket, a sample of households was asked what volume of the budget was allocated to the acquisition of each good and service that make up their final consumption expenditures. The sample was representative (35.998), the source was the ENIG of 2006-2007 and the periods were different, (daily, weekly, monthly, quarterly, and annually).

The already expanded and converted to annual expenditure information is the basis for the calculation of all the weighting structure that requires the construction of the index. This applies for each good or service within the basket, by income levels and cities.

The following expenditure groups by use functions and homogeneous in nature are obtained through an aggregation process of expenses associated to each good or service:

Food products, dwelling, apparel, health, education, recreation, transportation, communications, and other expenses, by income levels and, subsequently, the national weights.

Table 8. Weights of the CPI-98 expenditure groups

Description CPI	I Low	I Medium	I High	Total
Food products	41,42	30,18	14,70	29,51
Dwelling	28,31	29,12	31,37	29,41
Apparel	7,29	7,92	5,78	7,31
Health	4,26	4,06	3,39	3,96
Education	3,57	5,30	4,95	4,81
Entertainment	2,25	3,77	4,69	3,60
Transportation and communications	7,29	12,17	23,67	13,49
Other expenses	5,60	7,48	11,46	7,89

Source: DANE

Table 9. Weights of the CPI-08 expenditure groups

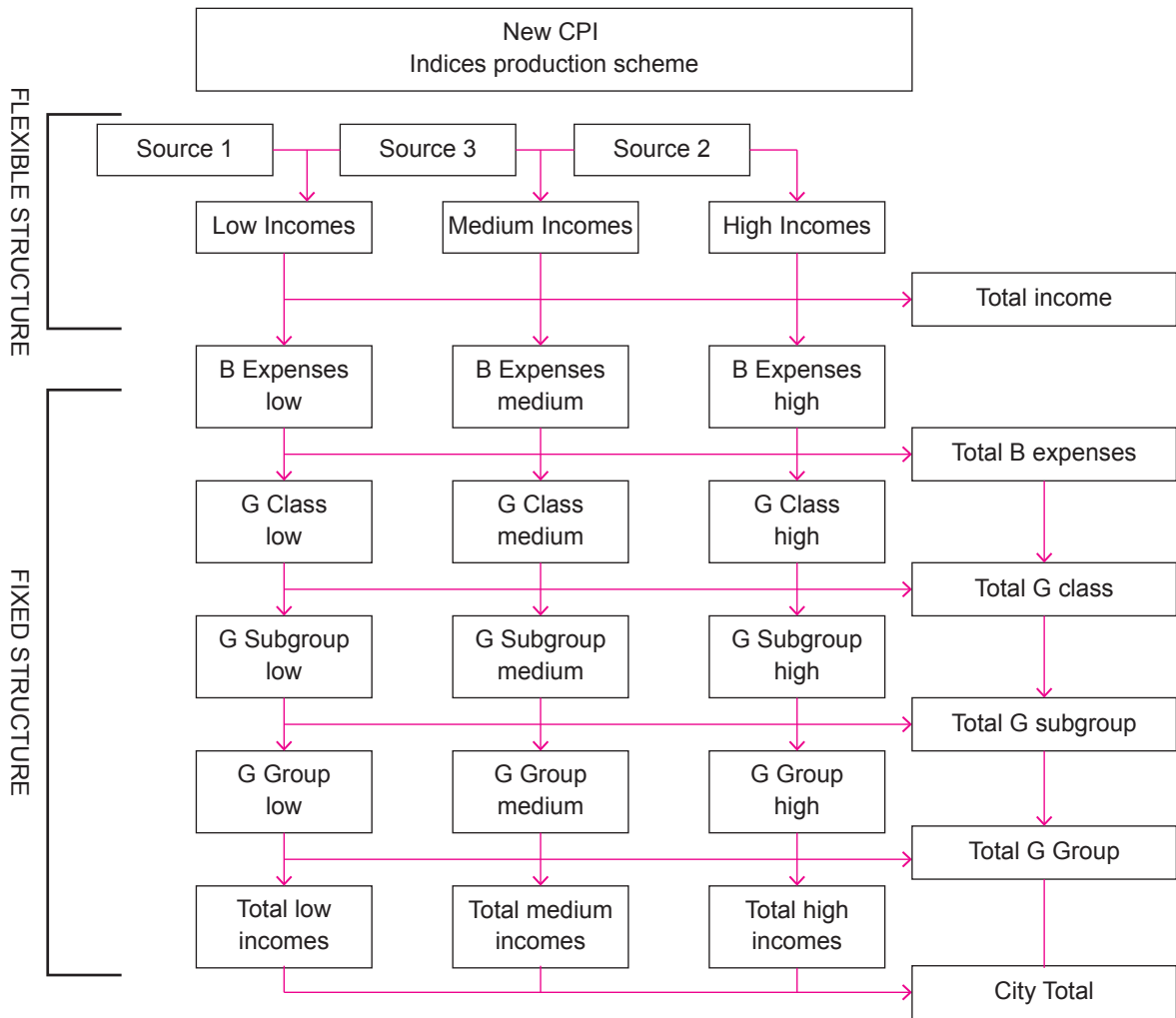
Description CPI	PI 08 (National Aggregate)			Total
	I Low	I Medium	I High	
Food products	34,66	27,09	18,24	28,21
Dwelling	29,74	30,42	29,66	30,10
Apparel	5,68	5,11	4,20	5,16
Health	2,04	2,41	3,39	2,43
Education	4,79	5,99	6,83	5,73
Entertainment	2,33	3,19	4,43	3,10
Transportation	11,03	15,32	23,88	15,19
Communications expenses	3,14	3,98	4,03	3,72
Other expenses	6,59	6,47	5,35	6,35

Source: DANE

Calculation Methodology of the CPI. The CPI calculation is based on the price relatives by reporting source to obtain the price relatives for each item by income levels, by means of geometric averages. Subsequently the weighted arithmetic average is used to obtain the indices for the basic expenditure by income levels and total. The aggregation by means of the arithmetic average enables to obtain the remaining structure indices. The total CPI calculation by city or national is based on the calculation of the indices by income level.

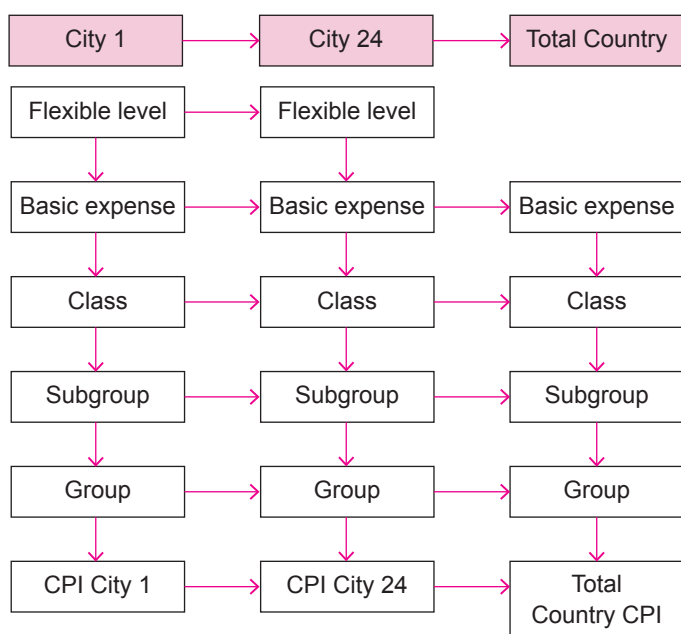
Production Indices

Figure 1. General, by income level for each city



Source: DANE

Figure 2. General for the National Total



Source: DANE

The calculation of indices by income level has the prices of the goods and services reported by the sources as base input. These sources are in turn classified as follows:

Large cities: in these cities their geographical location, the characteristics of the establishment, and the characteristics of the bulk of consumers who visit these establishments are considered for the classification of sources, by income level (source group assignment). According to this, the sources can be classified in three groups:

Sources in group 1, located in city areas where socioeconomic strata 1, 2 and 3 are predominant, that is, they are the purchase establishments of these strata because of their location in their area of influence.

Sources in group 2, located in city areas where socioeconomic strata 4, 5 and 6 are predominant, that is, they are the purchase establishments of these strata because of their location in their area of influence,

Sources in group 3, located in the cities' shopping areas, that is, where no socioeconomic stratum is predominant, or of residential use, and thus they constitute the purchase establishments for all socioeconomic strata.

Small cities: in the case of small-sized cities, since there is only one trade zone or shopping area located downtown of the cities, all trade sources are assigned the source group 3.

In the case of stratified-nature services such as utilities and rentals, sources should be classified as follows:

Group 1= low incomes, strata 1, 2.

Group 2= medium incomes, strata 3, 4.

Group 3= high incomes, strata 5, 6.

The source-group classification determines the level of income at which the relatives calculated by source participate:

- Sources in groups 1 and 3 participate in the calculation of the low-income level.
- Sources in groups 2 and 3 participate in the calculation of the medium-income level.
- Sources in group 2 participate in the calculation of the high-income level.

The procedure for the calculation of the different levels of the index is the following:

$$IRS = (P_t / P_{t-1})$$

- Simple relatives by source: represents the simple relative value of price behavior for a specification in a source, and its objective is to determine the behavior of prices for an item or variety between two time periods, on the same base unit for the same source. The calculation is:

$$IRS = (P_t / P_{t-1})$$

Simple relatives by source work with the following variables:

P_t = Price in the current period

P_{t-1} = Price in the former period

Their information sources are the Unique Collection Forms (FUR for its acronym in Spanish), in hardcopy and electronic means, consolidated in the database.

- Geometric average of simple indices by sources: the geometric average of the simple relatives by sources for each item or variety determines the average behavior of prices, for an item or a variety, among the sources that yield information. The form of calculation is:

$$PGISR = \sqrt[n]{ISR_1 * \dots * ISR_n}$$

The geometric average of simple indices by sources for each item or variety works with the following variables:

PGISR = Geometric average of simple indices, item or variety.

ISR = Simple index of relatives by source.

N = Number of sources which provides the price for the item or variety.

- Weighted average of geometric averages of simple indices: this one is for each item or variety; it determines the average price behavior of the items or varieties that constitute a basic expenditure. This calculation enables the update of the basic expenditure index of the previous period. The calculation formula is:

$$IPAGB = \frac{\sum \alpha PGISR}{\sum \alpha}$$

- The weighted average of the geometric averages of simple indices by item or variety for each basic expense works with the following variables:

IPAGB= Weighted arithmetic average index of items or varieties that make up the basic expenditure.

PGISR= Geometric average of simple indices by item or variety.

α = Weighting in relation to the basic expenditure.

- Local basic expenditure index: represents the index for the first level of the fixed portion of the CPI. Its purpose is to determine the level of the index number of the basic expenditure in the reference period, as a result of chaining the index of the former period to the weighted arithmetic average index of items or varieties that make up the basic expenditure. The calculation formula is:

$$IGB = IGB_{t-1} * IPAGB$$

The local basic expenditure index works with the following variables:

IGB= Basic expenditure index.

IPAGB= Weighted arithmetic average index of articles or varieties that make up the basic expenditure.

- Total basic expenditure index: represents the index number for the first level of the fixed portion of the CPI, for the aggregate of cities and each city. Its purpose is to determine the level of the index number of the total basic expenditure in the reference period, as a result of a double weighted average of the level of the basic expenditure indices by income levels. The form of calculation is:

$$IGBT = \frac{\sum \phi^* \beta * IGB}{\sum \phi^* \beta}$$

The total basic expenditure index works with the following variables:

IGBT= Total basic expenditure index.

IGB= Basic expenditure index by income levels.

ϕ = Weight or weighting of the national or local income level in the national total

β = Weight or weighting of the basic expenditure in the city's income level.

- Expenditure class Index: the expenditure class represents the index number for the second level of the fixed part of the IPC. Its purpose is to determine the level of the index number of the expenditure class in the reference period, as a result of a weighted average of the basic expenditure indices level that make up each class of expenditure. The calculation is:

$$ICG = \frac{\sum \beta * IGB}{\sum \beta}$$

The expenditure class index works with the following variables:

ICG= Expenditure class index.

IGBT= Basic expenditure index.

β = Weight or weighting of the basic expenditure.

- Expenditure subgroup index: the expenditure subgroup represents the index number for the third level of the fixed portion of the CPI. The form of calculation is:

$$ISG = \frac{\sum \beta * ICG}{\sum \beta}$$

The expenditure subgroup index works with the following variables

ISG= Expenditure subgroup index.

ICG= Expenditure class index.

β = Weight or weighting of the expenditure class.

- Expenditure group index: the expenditure group represents the index number for the fourth level of the fixed portion of the CPI. The form of calculation is:

$$IGG = \frac{\sum \beta * ISG}{\sum \beta}$$

The expenditure group index works with the following variables:

IGG= Expenditure group index.

ISG= Expenditure subgroup index.

β = Weight or weighting of the expenditure subgroup.

- Total index as an average of expenditure groups: represents the index number for the total CPI, as an average of expenditure groups. The form of calculation is:

$$ITG = \frac{\sum \beta * IGG}{\sum \beta}$$

The Total index as an average of expenditure groups works with the following variables:

ITG = Total expenditure groups index.

IGG= Expenditure group index.

β = Weight or weighting of the expenditure group.

- Total index as an average of basic expenditures: represents the index number for the total CPI, as an average of basic expenditures. The form of calculation is:

$$ITG = \frac{\sum \beta * IGB}{\sum \beta}$$

The total index as an average of basic expenditures works with the following variables:

ITG= Total expenditure groups index.

IGB= Basic expenditure index.

β = Weight or weighting of the basic expenditure.

Special calculations:

Special calculations are explicit-quality adjustments. A different calculation is made for each item:

Vehicles: the calculation is performed using a components' model, that is, the vehicle's most important parts and characteristics are identified along with any changes that occur.

Computers and printers: the calculation is performed using a hedonic model, provided by the Agency of Statistics Canada, which enables the determination of physical differences, such as hard-drive capacity, processing speed, RAM, monitor type, and computer brand.

Lotteries: the calculation is performed using a hedonic model, by means of which the modified price variation in the current period and in the former period adjusted by the probability of winning is incorporated into the index.

Indicators for dissemination

The main indicator generated by the CPI is the index; however, there are *variations, contributions, and participations* (monthly, year to date, and twelve months) that are also obtained from the investigation.

- **Monthly variation:** this is the relation of the index in the reference month (li,t) with the previous month's index (li-1,t) , minus 1 times 100.

$$VM = \left(\frac{\text{Reference month index}}{\text{Previous month index}} - 1 \right) * 100$$

- **Year-to-date variation:** this is the relation of the index of the reference month (li,t) with the index in the month of December of the previous year (id, t-1), minus 1 times 100.

$$VAC = \left(\frac{\text{Reference month index}}{\text{December of the previous year index}} - 1 \right) * 100$$

- **Twelve months variation:** This is the relation of the index in the reference month (li,t) with the index of the same month in the previous year (li, t-1) minus 1 times 100.

$$V12M = \left(\frac{\text{Reference month index}}{\text{Same month previous year index}} - 1 \right) * 100$$

The contribution enables the measurement of the contribution in percentage points of each hierarchical level (expenditure group; expenditure subgroup, expenditure class, and basic expenditure) to the month variation, year-to-date, and twelve months, of the CPI total.

- **Monthly contribution:**

$$CM = \left(\frac{I_a \text{ previous month}}{TN \text{ index previous month}} \right) * P_a * VM_a / 100$$

- **Year-to-date contribution:**

$$CAC = \left(\frac{I_a \text{ December previous year}}{TN \text{ index December previous year}} \right) * P_a * VAC_a / 100$$

- **Twelve months contribution:**

$$C12M = \left(\frac{I_a \text{ Same month previous year}}{TN \text{ index same month previous year}} \right) * P_a * V12M_a / 100$$

Where:

a: Hierarchical level.

I: Hierarchical level index.

TN: National Total.

P: Weighting of the Hierarchical level.

VM: Monthly variation of the hierarchical level.

VAC: Year to date variation of the hierarchical level.

V12M: Twelve-months variation of the hierarchical level.

The *participation* is the percentage of the contribution explanation of each basic expenditure, expenditure class, subgroup, and expenditure group in the variation, of the total index.

- **Monthly participation:**

$$PM = \left(\frac{CM \text{ hierarchical level}}{\text{Monthly variation}} \right) * 100$$

- **Year-to-date participation:**

$$PAC = \left(\frac{CAC \text{ hierarchical level}}{\text{Year to date variation}} \right) * 100$$

- Twelve-months' participation:

$$P12M = \left(\frac{\text{C12M hierarchical level}}{\text{twelve months variation}} \right) * 100$$

Linking methodology: in order to provide statistical information users with the necessary tools to make use of the CPI information adjusted to their needs, the way of linking the CPI-08 series with the old CPI-98 base is indicated below.

The linking procedure, also known as base change, is obtained from applying a rule of three (3), where the *December 2008 index base 98* (whose value for this example is 191,63) is multiplied by the index published in *January 2009 base 08* (e.g. 100,59). This result should be divided by one hundred (100). The procedure for the subsequent months consists of changing the value 100.59 by the index at which the change of base is needed only.

The calculation formula is the following:

$$IPC_{\text{Base98}}^{\text{Jan09}} = \frac{(I_{\text{Base98}}^{\text{Dec08}} * I_{\text{Base08}}^{\text{Jan09}})}{I_{\text{Base98}}^{\text{Dec08}}} = \frac{(191,63 * 100,59)}{100} = 192,76$$

In the event that the previous information is needed, expressed in the new base, the old values should be brought to the current base. Thus, if the value of November 2008 Base 98 is needed, in case it has been calculated under the current base, the reference index (in this case November 2008, which value is 190,78) should be multiplied by 100 and then divided by the December 2008 index of the previous base (which value is 191,63). In order to obtain all the previous data with the new base, only the reference period is changed.

The procedure that must be used to link the series depends, in each case, on the information needs. Finally, it is important to point out that none of the calculated variations will be modified because of the linking procedure.

An alternative method consists in obtaining a linking coefficient as the ratio between the last index of the previous base (in this case, December 2008 CPI base 1998 =191,63) and 100, and then multiplying each index of the new base by this coefficient. The results obtained by this alternative method are identical to the method explained above.

Sampling design

Sampling frame

Education: the 2006 Directory of Schools and the 2007 Directory of Higher Education Institutions.

Rented dwellings: the geographic base at the level of blocks, corresponding to the 2005 General Census.

Other goods and services: these types of sources are selected by non-probabilistic methods.

Sample design

The sample design for the calculation of indices is a non-probabilistic design; however, a sample size is controlled which guarantees a minimum number of sources per item given the change in prices.

This process of controlling the sample size is performed monthly, for the prices of each item, with information on source and city.

The specific procedure for each city and each month is the following:

Step one: the geometric average of the variation between the current month and the previous month is calculated for each item.

Step two: the variance of the items' relative price indices is obtained with the geometric average.

Step three: the quotient between the square root of the variance and the estimated value of the price (relative error) is calculated.

Starting from the fact that there is seasonality in the prices of items, the above steps are carried out for each month of the year, getting 12 price-relative errors per item.

- a. The record with the maximum relative error for each item is identified.
- b. From this record, the variance and relative price index information is taken.
- c. Finally, the sample size is calculated with the variance and the relative price index (5), considering a sampling error of 5 %.

On the other hand, for the generation of errors that could be evaluated (relevant) it is established that the analyzed item should be composed of at least of 5 sources.

According to the above, the generalities of the sampling design for some research groups that make up the CPI are as follows:

Education: the design is a stratified random sampling with a selection probability in proportion to the size of the institutions, in terms of number of students. The stratification variables are: city, sector (public, private), and school teaching level (preschool, elementary school, middle school, and high school).

Rented dwellings: a probabilistic sampling design of conglomerates, and stratified.

- Conglomerates: constituted by the existing rented dwellings in the selected blocks.
- Strata: city, socioeconomic stratum, and number of dwellings on the block.

Other goods and services: these sources are selected by non-probabilistic methods. The selection criteria of the sources are focused on representing the different establishments where consumers purchase their goods and services.

Sample size

Sample sizes are calculated in function of the variability of the geometric average and a fixed relative error (E_{rel}) of 5 %, taking into account the following formula:

Where:

$$VAR_i [PG] = \frac{\sum_{j=1}^n (PG_{ij} - PG_i)^2}{n-1}$$

Selection method

Education: the institutions selected in each stratum are the ones with a probability in proportion to the quantity of registered students.

Rented dwellings: housing blocks are classified by stratum, and subsequently blocks are selected at random. Rented dwellings within each block are listed and the rental rate is inquired.

Other goods and services: the selection of establishments is carried out based on the geo-referenced information and with the help of expert staff that know the places where the different goods and services are usually purchased.

2.2.2. Design of processing and imputation methods. The imputation procedure begins with the on-site detection of the temporary absence of the specification at the source that is being visited. In that case, the technical novelty *Waiting Period* (PE for its acronym in Spanish) is registered in the Data Capture Device (DMC for its acronym in Spanish), or in the Unique Collection Form (FUR) – in hardcopy.

The imputation process is carried out with all the information collected for the month in the city, at item and income levels.

The information processing system groups the information rendering sources, according to the source group of the same, that is, group 1, 2, and 3.

The system proceeds to calculate the geometric average of relatives by sources of each item or variety, according to the following general formula, controlling by source group:

$$PGR = \sqrt[n]{Pt/Pt - 1 * \dots * Pt/Pt - 1}$$

Where:

PGR= Geometric average,

Pt= Current price,

Pt-1= Previous price.

Subsequently, this geometric average of price relatives is applied to the previous price of the item or variety that has the waiting period marked as follows:

$$PA = PANT * PGR$$

Where:

PA= Current price.

PANT= Previous price.

PGR= Geometric average.

This procedure does not alter the actual variation observed in the sources, where the item or varieties, which recorded a price change, were available. It is important to note that this imputation procedure is subject to the application of the following norms of the waiting period: First, not more than one waiting period can be applied consecutively for the items or varieties that are part of the index basket. Second, all waiting periods used in a collection period should not exceed 5 % of the total quotes for a specific item.

2.2.3. Design of instruments. To collect the prices of the CPI basket of goods and services (CPI-98), the *Unique Collection Form* (FUR) was designed, which in the early years of the research was filled-out manually. Since August 2004, DANE has been implementing the use of *Data Capture Devices* (DMC) to capture the prices directly at the visited sources. This is the technology that supports the work on-site in the CPI-08. For some specific items and in areas of danger of theft of the DMCs, the form is filled out manually.

The design of the FUR is preserved in the *SatForms* software application installed in the DMCs.

General specifications about the unique collection form

DANE logo appears in the first module.

In this first module the identification of the research is indicated: Consumer Price Index, CPI-08.

In the first module, section one is indicated (1-FUR Number): the system assigns a consecutive number to the source when the form is previously filled out.

The *law of statistical reserve* is stated in the FUR, which informs the sources about the confidentiality with which DANE's information is dealt with.

In the second module, section two appears (2-City): to include the city code in the first space, according to the Political-Administrative Division (DIVIPOLA), and in the second space, the name of the city.

In the second module, under Section Three (3-Source) the source code, made up of 20 digits, is registered, within which the geo-references of the establishment namely, the sector, section, block, group, class, and source code, are indicated (see table 10):

Table 10. Instructions for completing the fur

Field description	Length	Initial position	Final position
Department code	2	1	2
Town code	3	3	5
Sector number	4	6	9
Section number	2	10	11
Block number	2	12	13
Group code	1	14	14
Class code	2	15	16
Source consecutive number	4	17	20
Total	20		

Source: DANE

In the second module the following information is also included:

- Name or business name of the establishment.
- The city area to which it is assigned for collection.
- The establishment's address.
- The establishment's phone number or numbers

In the *second module* section 4 (4-PERIOD) appears. There, the identification of the period to which the information relates to is included. The year is written as a four (4)-digit number, the month is written as a two (2)-digit number, and the delivery as a one (1)-digit number.

The digits to identify the months are:

January 01, February 02, March 03... December 12.

In this second module there is a Table with information of items previously filled out per source, with the purpose of controlling the number of items or price quotes that are taken at each source.

Body of the form

In section 5 (5-CODE) the item code is registered, which is made-up of seven (7) digits corresponding to the group, subgroup, expenditure class, and basic expenditure, in which the item is classified.

Section six (6- SPECIFICATIONS) - The name of the item and its description or its detailed specifications are registered.

The concepts that identify the item, the specification are particular by source, since each establishment markets a particular brand, a reference, a variety, a model, etc.

Section seven (7-BASE UNIT) - The amount and unit that correspond to the unit of measure that the system internally has for the item, are indicated.

The system converts prices into this unit whenever the collected unit (column 15) is different from the base unit.

Section eight (8- PREVIOUS UNIT) – This refers to the unit of measure or quantity taken in the previous collection.

Section nine (9 - PREVIOUS PRICE) – The price obtained (without any conversion) in the previous collection is indicated. This price corresponds to the cash price and with VAT, as applicable.

In section ten, Waiting Period (10-P.E), a cross (x) is registered when the item is granted a waiting period without information that has been methodologically established.

Section eleven - Immediate Substitution (11-S.I) – This is used to indicate (by means of a cross (x)) if the immediate substitution of the item's specification was performed. This novelty applies when one of the item's secondary characteristics changes and it is registered in the index as an effective collection of prices of the item involved in the calculation.

Section twelve, Complementary Source (12-F.C) – This section is filled only when the collected information comes from a complementary source, by registering a cross (x) and completing the information requested in section 19.

Section thirteen - Still Among the Best-selling Products (13 C.M.V) – This is marked with a cross (x) if the item is still among the best-selling articles.

Section fourteen - Reference Change (14-C.R) – Marked by a cross (x) when the item modifies its primary specifications, as long as it does not belong to the perishable foods' group.

In section fifteen, Collected Unit (15-U.C) – The base unit of the item is expressed in terms of its contents, that is, in units of weight (grams), units of volume (cubic centimeters or ounces), or in units of longitude (meters). In this column no decimal places or letters should be entered.

On the contrary, if the base unit of the item is expressed in terms of Units (1, 2, 6, etc.), in section fifteen (15-U.C), for example in the case of medicines, the number of pills, tablets, or capsules (1, 4, 80, etc.) should be recorded, or, for most articles simply one (1).

If the base unit is one (1) and it is an item in liquid or solid presentation (not tablets, pills, capsules, diapers, or condoms), in section fifteen (15-U.C) the number one (1) should be recorded. For this case, the content of the item in cubic centimeters or in grams will be indicated in section six (6 NAME – ITEM – SPECIFICATIONS).

In section sixteen (16 – CURRENT PRICE), the price obtained for the item in the current period is recorded. This price corresponds to the cash price and with the Value Added Tax (VAT).

In section seventeen, Observations (17-OBS), the number assigned to each one of the established observations is recorded.

In the event of a non-encoded observation, the explanations related to novelties that took place in the collection for the different items are recorded in the back of the form, or in the field established for this purpose in the DMC.

Section eighteen (18-COMPLEMENTARY SOURCES) is used to identify the establishments selected as complementary sources when it was not possible to continue taking information for one or more items in the original source. This literal is associated with how section twelve (12-F.C.) is completed.

The collection date must be filled out in the following format: dd/mm/yyyy (day with two digits; month with two digits; and year with four digits). The same format must be applied when filling out the Date of Inspection.

The fields to indicate the responsibility of the collector, the supervisor, and of the source should be filled out by each responsible person. Both the collector and the supervisor must write their names and signatures, as well as the signature or stamp of the informant source.

Statistical production

The statistical production is developed in two levels: the Local level and the Central level. Each of DANE's six Territorial Branches, with which the entity counts to advance the established processes in the functional model, is included in the Local level. These divisions cover the information on the twenty-four selected cities in the country. Moreover, the central level or *DANE Central* refers to statistical production processes that take place in the entity's headquarters in the city of Bogotá.

At Local level, each of DANE's offices and sub-offices organizes, prepares, collects, conducts, monitors, and captures information related to the investigation. Each Territorial Branch is responsible for the quality of statistics, which are generated in relation to the monthly collection and analysis procedures; therefore, DANE demands an accurate and effective fulfillment and compliance from its staff and personnel in charge of coordinating and supervising the process in the different stages.

At the *DANE* Central the statistical production is related to in-site logistics of activities, ranging from production to the analysis of results.

3

PRELIMINARY ACTIVITIES

3.1. AWARENESS-RAISING PROCESS

The awareness-raising process is performed by the *collector* at the time he/she visits the different sources, informing them of the objectives, purposes, and benefits obtained from the information provided to DANE, thus, creating statistical awareness in each of the informants. Likewise, at the time of the visit the *collector* provides the source with a brochure that contains complete information on the research.

3.2. STAFF HIRING

Once the *DANE Central* sends the resources and required profiles for personnel selection to the *Territorial Branches* the divisions then search for candidates that meet the requirements and comply with the current regulations for this exercise. This process is based on verifying and classifying the suitability of each candidate that applies for a position.

The positions of the operations scheme of the CPI that are performed within the Territorial Branches by personnel in charge of contacting the informant sources and collecting information on prices are:

- *Collectors*: personnel in charge of collecting the information related to the surveys.
- *Analysts*: personnel in charge of control functions, operative coverage, and management analysis.
- *Supervisors*: personnel in charge of validating the information delivered by collectors.

Similarly, the positions at the *DANE Central* are those related to thematic, logistics, and systems activities.

- *Experts*: personnel in charge of defining the methodological process of the research and of providing the guidelines to the logistics and IT personnel.
- *Logistics Staff*: this is the personnel in charge of the analysis, revision, correction, and filtering of the quality of the information received from the Territorial Branches.
- *Systems staff*: personnel in charge of support, maintenance, and development of database applications, that results from the research.

Once the personnel is hired a training is carried out in groups, using tools such as video projector, brochures and guides. An overview of the research is provided referring to aspects such as number of updates, coverage, objectives, importance, and methodology.

This activity makes emphasis on each working team according to the position; group of collectors, supervisors, analysts, and support staff, informing these teams about all the processes, so that they can have information on the investigation and therefore make their contributions at a given time. In addition, practical workshops are conducted presenting field case-studies and their respective solutions, in order to strengthen the staff's knowledge of the investigation.

3.3. COLLECTION OF INFORMATION

3.3.1. Operative organization Structure. At the Central Level, the functional structure counts on a technical coordinator in charge of technical and methodological aspects, a logistics team coordinator with support for the research, and professionals who are responsible for the analysis, validation, filtering, refinement and flesh out of the information reported by each of the cities that conform the geographical coverage of the research.

At territorial level, the functional structure has an operations coordinator, a coordinator of indexes, support for the research, analysts, supervisors and collectors or interviewers, depending on the number of item quotations that are followed-up in each city.

The *operations coordination* is responsible for the research within the Territorial Branch.

The *indices coordinator* is responsible for the operational process and for its proper functioning.

The *Technical support* is responsible for accompanying the coordinator in all the coordination, instruction, training, and monitoring tasks assigned by DANE Central, or by the indices coordinator.

The *analyst* is responsible for analyzing and filtering all the collected information, so that it is of an optimum quality when delivered to DANE Central, according to the time schedule that has been established.

Supervisors are responsible for validating the information collected in the field and for making the necessary changes, in order that the resulting information is reliable and of quality.

Collectors are responsible for visiting all sources that are part of the research sample; they collect the information and, at the same time, develop the awareness-raising process with each of the informants.

3.3.2. Operational scheme, method and procedure for collection. The CPI data collection is based on the completion of the FUR. Currently this form is completed with a mixed collection system: on paper and on the *Data Capture Device* (DMC). The price collection is carried out during a personal visit, information over the phone should not be taken, and in no case is the interview to be delegated to another person different from the assigned collector, or shown to third parties.

The supervisor validates the collected information and visits the source one or two days after collection, in order to verify that the information was properly collected.

The analyst, as mentioned above, is in charge of analyzing and refining all the information that was collected.

Working teams are formed in a ratio of three (3) collectors per one (1) supervisor and one (1) analyst.

The price collection process is an exercise carried out by DANE's local offices, thanks to a team of collectors, who visit directly the establishments that offer goods and services that are contained in the basket.

A routes system is used to distribute the work. This system consists of the programming of daily routes jointly by collectors and supervisors, once the monthly information is completed.. The routes are evenly distributed and daily visits are programmed on collection templates, indicating the sources to be visited during the month. Daily, the analyst loads the DMC with the programmed information and, at the same time, verifies that the schedule is met, according to what has been established in the program. With the DMCs loaded, collectors go to the field and conduct interviews to each of the respondents, thus collecting the required information.

The price collection is determined based on a description of specifications associated with each good or service; these specifications include brands, presentations, units, observations, and more features on what the consumer purchases the most. The sources from which the quotations are taken correspond to the establishments referred by consumers as the most-frequent purchase sites. These establishments are included in ten categories, which are listed below:

Table 11. Category of establishments where more goods and services are purchased

Type code	Informant source type
1	Posts in marketplaces, including mobile stands
2	Supermarkets and private stores
3	Compensation or allowance funds, cooperatives, employee funds and commissariats
4	Non-specialized corner stores
5	Stores or specialized shops
6	Drug stores, apothecaries or pharmacies and perfume stores
7	Specialized establishments – service providers
8	Restaurants or outlets of foods prepared in-chain
9	Other establishments, including rented dwellings
10	Hypermarkets

Source: DANE

The collection process also considers the frequency with which prices of the items included in the CPI tracking basket change. Hence, for those items with a high-rate of change in prices (e.g. food), the process is carried out on a monthly basis.

Other items are collected with a different periodicity, because the frequency of price changes is lower. For example, personal care items, household-cleaning items and medicines are collected bimonthly.

There are quarterly collection periodicities. This is the case for household appliances, apparel, etc. In the case of rented dwellings, the frequency of price collection is every four months. Prices for educational services are collected every six months and annually.

Based on this classification, the CPI has determined monthly sample fractions for all goods and services. Thus, there is always a representative price, to capture “unexpected price” increases in any of the goods and services in the tracking basket.

Additionally, it has been determined that some goods and services, whose price increases from time to time because of an administrative act, must have a continuous price-tracking, and report only when an effective increase is present, e.g. cinema, lottery, urban transport, fuel, etc.

The data collector will receive information concerning the working or collection route of the different goods and services, from the technical assistant, or from the coordinator.

The collector must personally visit the informants and directly obtain the retail prices of the goods and services. In the event that it is not possible to interview the informant, the collector has to report this situation to the supervisor, who should indicate other means to obtain the data, always leaving a written record of the situation with an observation or comment on the collection form.

If during the collection visit the informant is very busy, the collector should set up another appointment at a later time or day that suits the informant, in order to obtain consistent outlet prices.

3.3.3. Mapping used. The CPI uses mapping or cartography for the selection of dwellings for rent, based on a probabilistic sample. For this, the properties for rent are listed and then, cartography of planes of the sectors and a 1:5000 scale map are used.

3.3.4. Supervision and control of the operation of collection. The process of supervision involves a series of activities designed to control the quality of the collected data, to detect particular situations with informant sources, and special or atypical behavior of prices obtained in the collection.

The activities of the supervisor begin with the delivery of the list of sources to be collected monthly to the assigned collector. Sources are classified by collection areas in order to verify that latter list corresponds to the sources to be visited in the corresponding period. The delivery of the list could be either in hard copy or in magnetic means.

Once all the information is collected in accordance with the daily analysis report and route selection per supervisor, the source(s) is (are) visited, to confirm that the information provided is true; and furthermore, the guidelines outlined by the coordinator and included in the *Filing-out Manual* are also applied in the process. The supervised Items (codes) should be entered in the supervision template and the guidelines for their registration deal with representation of sources, selective supervision, rises, falls, and technical novelties.

As for inconsistencies found in the supervision process, corrections are made, relevant comments are noted and the information is returned to the analyst.

Every month the local coordinator and the professional assistant will define the number of sources to be supervised, taking into account the different areas and collectors, as well as indicating the program schedule and the supervision results on a report.

In the supervision process it is important to bear in mind two fundamental aspects: The source data and the data on items or varieties:

Source data: in regards to the source, the information must be verified on the field with respect to the following items:

- Name.
- Address and phone number.
- Collection period.
- Group and class.

When confirming or correcting the data, any novelty should be informed to the analyst in charge according to the situation that arises

The correct classification of group and source class should be verified, in a new source.

Data on items or varieties: In this aspect it is necessary to check that the information on each item is complete and contains the minimum specifications:

- Brand or variety.
- Primary characteristics.
- Secondary characteristics.
- Base unit.
- Previous unit and previous price.
- Collected unit and current price.
- Technical novelties, if required.
- Comments or observations, if required

The product information is reviewed after checking the information from the source. Accordingly, it is necessary to verify that the monitored items have the minimum specifications to identify the item or, otherwise, add what is deemed necessary to recognize the variety that is being compared between the quoted periods.

The supervisor should perform calculations of price variations taking into account the collected unit, the current price and the observations of the collected period in relation to the former period. For every difference of information that exists between the two periods that are being compared (whether an absence or a change), an applied technical novelty should be used to explain it.

Doubts or omissions are clarified and any inconsistencies encountered in the process are corrected by interacting with collectors, these doubts or omissions are recorded in the supervision template or, if subject to conciliation, they are discussed with all the operative team to implement the novelty that adjusts the best, according to the CPI methodology.

If the data provided by the source is inconsistent with respect to that obtained by the collector, it is important to report this novelty to determine whether the informant is the same or not, and if so, check the information provided by the source jointly with the supervisor,. Potential biases in the information can be detected in this manner, in case there was a different informant for each visit that was carried out, thus improving the quality of the data obtained.

The supervisor has to schedule accompaniment for every collector in each working route, in order to evaluate the presentation, identification, expression, interviewing technique (access to information), polling technique (to complete or clarify responses) and others that guarantee the quality of the information collected in the field.

Monthly, the collection of prices will undergo a supervision process led by the local coordinator, which purpose is to verify facts such as:

- Existence of the informant source.

- Volume and quality of the information collected, with verification of the collection date.
- The treatment given by the collector to the informant and vice versa.
- Other aspects of the collection process.

Finally, it is important to mention that the role of the supervisor in matters related to the project's quality control is decisive, since he is the person in charge of validating the information collected in the field and make the parameters established for quality control available.

3.3.5. Classification and management of surveys or records. Once the information is collected, each DMC is synchronized with the computer and the loading and unloading of information from the digital format is performed automatically, prior sorting of sources by group-source, class-source and by area. Subsequently, the collection of sources and items is monitored by analyzing each data source and by considering quality and reliability indicators.

3.4. CODIFICATION AND ANALYSIS PRODEDURE

The analysis process is directly linked to the quality of the collected information. In the analysis each of the fields included in the FUR is verified and examined and the information is refined. In addition, the forms that have inconsistencies and significant variations are selected for on-the-spot supervision, separating them from those whose information has been correctly and completely refined.

Given that the information is systematized, it is possible to arrange it in the system for its verification by sources, items, variations and technical novelties according to the analysis needs.

3.5. CAPTURE AND CONSOLIDATION OF DATA

Data are consolidated with the download of information from the DMCs to the application to feed the database research.

The DMC is an instrument that enables the collection of information on one or several items that exist in a source, which may have, different specifications for each of the cities.

Once an area is assigned to each collector and supervisor the following fields can be accessed through the DMC:

- Password to log into the application of the device.
- Names of the scheduled sources.
- Header of the form: general data of the source: name, code, address, phone number, area and month of collection.
- Item code: specification, unit and price, technical novelties and observations or comments.

3.5.1. Consolidation of files. In the annual program the means of sending the information (magnetic or physical) is established and the scheduled dates for the receipt and consolidation of such information by the DANE Central are identified. According to the program, the office or the sub-office must generate a magnetic file for each set of information scheduled for sending in the month, through the decentralized module. These files are

sent by email to DANE Central. The information of the month for the CPI research is received on a weekly basis, and this information is loaded directly into the centralized application developed for the research.

In the loading process (receipt and consolidation) the system can identify files that contain inconsistencies, the IT systems' area checks and corrects these situations in the city.

Once the receipt and consolidation of information for each of the 24 cities is done, the process of analysis by DANE Central takes place. For this the workloads are assigned to each of the professionals who are part of the logistics production team. The analysis includes the validation of information, the correct application of technical novelties, the analysis of average prices, the verification of qualities and specifications of each input, the detection of inconsistencies and the request of novelties by supervisor.

3.6. PROCESSING AND IMPUTATION

The general imputation method for all price and cost indices pretends to capture the price effect that the indicator must reflect when the collector faces the absence or lack of an item or variety that he/she wishes to collect.

In the specific case of the CPI in the IT application, for calculation, the method of price imputation has been programmed, for the records that present the technical Waiting Period (PE) novelty, according to the method agreed, and as explained in the processing design and imputation section. The margin established for the application of the technical novelty PE is controlled by means of quality and reliability indicators.

3.6.1. Coding method. The Thematic and Logistics teams have developed a table with the possible observations to the items that compose the basket; this table has been included in the capture applications. Every specification of items is codified and given a consecutive number as soon as it is captured, that is, an available consecutive number is given to each specification as soon as a new one is registered in the table. Also, the thematic team has defined the codification of items for the calculation process. As regards the sources the codification is a consecutive number by city. The codification of the technical novelties is presented below:

1 = Informs normally.

2 = Change of reference.

3 = Waiting period of DANE Central.

4 = Waiting period of the city.

5 = Quote exits.

6 = New quote

3.6.2. Adaptation of files. Generally speaking, the capture, processing and exit variables are the following: Sources, items, prices, technical novelties, observations, specifications, indices, variations of prices, and variations of indices. Each of these variables presents month-to-month values, which are updated in an orderly manner, according to the different processes in which they are involved. The historical record of each of the capture variables is located within a database; such information is updated over time.

The database is implemented in *Visual FoxPro*, with each city having a database. Weekly, the files of transfers are consolidated in *Visual FoxPro* in DANE Central.

3.6.3. Data editing. The module of information capture is located in the decentralized module used by analysts in the cities and serves to analyze or to incorporate the information that is not being captured with the DMCs. The importance lies in the fact that it enables the identification of how many effective records have been collected and how many of them are missing from the universe (Figure 3).

Also, it counts on the module of reports that lists the available reports, to follow-up on the sample. These reports can be generated on screen, in MS Excel (.xlsx) or MS Notepad (.txt), and can be seen in Figure 4.

Figure 3. CPI capture form

Código	Artículo - Especificaciones	Unidad Base	Unid. Ant.	Precio Anterior	P.	E.	S.	F.	C.	M.	V.	C.	R.	Unid. Rec.	Precio Actual	Obs.
6210101	TEXTOS PRIMARIA	1Urd.	1	126,200.00	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0	0.00	
6210103	TEXTOS SECUNDARIA MEDIA	1Urd.	1	133,260.00	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0	0.00	

Source: CPI application, DANE

Figure 4. Reports generated in the CPI application

Reportes

- Artículos que entran en periodo de espera
- Artículos con sustitución inmediata
- Artículos con cambio de referencia
- Artículos cotizados en fuente complementaria
- Artículos que no continúan entre los más vendidos
- Variación de precios por artículo
- Artículos cuya variación excede el rango esperado
- Artículos sin precio y sin periodo de espera
- Resumen de Novedades Técnicas

Opciones

Source: CPI application, DANE

Centralized module: used by DANE Central's analysts to verify the consistency of the information of the cities. In addition it counts with several system reports that enable the control of coverage and quality of the information.

Analysis environment: the necessary tools to carry out an exhaustive refinement of the information contained in all the reports sent by each of the cities of the indicator are found in this module. Here the information required can be visualized according to data sending, the city, the article and the selected ranges of variation (Figure 5).

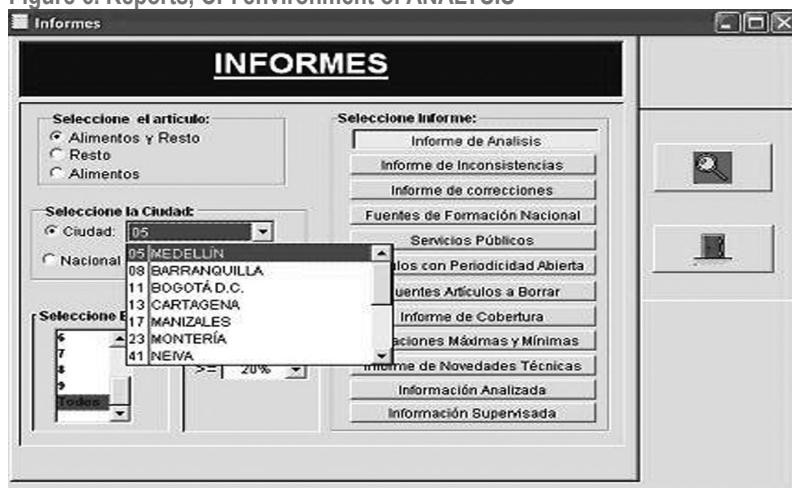
The module of reports enables to obtain information according to data sending,, the city and the selected group of expense (Figure 6).

Figure 5. CPI Environment of ANALYSIS

Información de la Fuente		Periodo Anterior			Periodo Actual			Novedades Técnicas				Análisis													
\$\$\$	Cltu	Env	Grp	Códi.	Nombre	Cantid.	Precio	\$Base Ant.	Mk	Cantid.	Precio	\$Base Act.	Mk	CM	Varia.	PE	CR	SI	FC	CMV	Obs	As	Sup	Análisis	
05	1	1	0029		SUPERMERCADO ESTELAR	500	790	' 790	1	500	790	' 790	1												✓
05	1	1	0108		PRODUCTOS CALOMA	500	850	' 850	1	500	850	' 850	1												✓
05	1	1	0114		SUPERMERCADO ORIENTAL	500	800	' 800	1	500	800	' 800	1												✓
05	1	1	0480		MERQUEFACIL SAN JAVIER	500	890	' 890	1	500	890	' 890	1												✓
05	1	1	0523		ECOMERCA	500	750	' 750	1	500	750	' 750	1												✓
05	1	1	1398		CONSUMO CASTILLA	500	870	' 870	1	500	870	' 870	1												✓
05	1	1	1465		MERCADOS PEDREGAL	500	850	' 850	1	500	850	' 850	1												✓
05	1	1	1553		AUTOSERVICIO LA 98	500	750	' 750	1	500	750	' 750	1												✓
05	1	1	1781		MERCADO VELEZ-EDILBER	500	830	' 830	1	500	830	' 830	1												✓
05	1	2	0003		ALMACEN EXITO POBLADO	1,000	1,780	' 890	1	1,000	1,870	' 935	1		5.06							4			✓
05	1	2	0011		POMONA POBLADO	3,000	5,550	' 925	1	3,000	5,720	' 953	1		3.06							4			✓
05	1	2	0070		MERCADOS AMARU	500	850	' 850	1	500	850	' 850	1												✓
05	1	2	0129		POMONA DE OVIEDO	500	950	' 950	1	500	980	' 980	1		3.16							4	✓	✓	✓
05	1	2	0372		SUPERMERCADO OLIMPICO	500	800	' 800	1	500	800	' 800	1												✓
05	1	2	0797		MERCOOP S.A.	500	850	' 850	1	500	900	' 900	1		5.88							4			✓
05	1	2	3265		SUPERMERCADO EL GRAN	500	850	' 850	1	500	850	' 850	1												✓

Source: CPI application, DANE

Figure 6. Reports, CPI environment of ANALYSIS



Source: CPI application, DANE

3.6.4. Imputation criteria. In the case of the CPI, there is only one imputation method, which is applied solely for cases of the technical novelty PE. This method is applied during the temporary absence of the price for the collection month.

The index calculation application uses the formula to determine an imputed price every month. With this imputed price the variation for the next month is determined, in which there will already be information; if this is not the case, and the technical novelty to be applied will be Quote exits (IS)¹⁰.

There is only one imputation level and it is for the temporary absence of a price in the collection month. This process is performed controlling the following variables:

- The system of information processing groups the sources that provide information according to their source group, that is, groups 1, 2 and 3.
- The system calculates the geometric average of relatives by sources of each item or variety, according to the following general formula, controlling by source group:

$$PGR = \sqrt[n]{Pt/Pt - 1 * ... * Pt/Pt - 1}$$

Where:

PGR= Geometric average

Pt= Current price

Pt-1= Previous price

- The system proceeds to apply this geometric average of relatives to the previous price of the item or variety that has the PE marked.

$$PA = PANT * PGR$$

Where:

PA= Current Price

PANT= Previous Price

PGR= Geometric average

3.6.5. Data integration. In order to come to use the micro data there are several processes that are to be completed before:

First, in-the-field capturing or capturing of the format in DANE's offices

Second, synchronization of this mobile capture through the synchronization application in each of the index's cities. (Each city does this separately).

Third, decentralized or by-city data revision and supervision. Whether at decentralized or by-city level each must review its data (analyze and supervise them, as necessary).

Fourth, weekly email sending of information to DANE Central in compressed files generated by the application.

¹⁰ Quote exits is a technical novelty that is used to face the definitive absence of the quote of an item or variety.

Fifth, reception and consolidation by DANE Central, through its application for file sending.

The *micro-data* is obtained after following the previous steps, in order to proceed to the central analysis.

4 ANALYSIS

4.1. STATISTICAL ANALYSIS

Once the process of reception and consolidation of the information sent by the cities has been completed, each analyst is responsible for carrying out the analysis of the information on the assigned cities.

An analysis module is available for the CPI analysis of results. The fundamental objective of this module is to evaluate the consistency of the information collected in each city, in order to detect errors and correct prices or new events of each sending that took place during the confirmation with the cities.

This process takes into account the assessment of data consigned for the reference period in each of the articles that conform the basket, an assessment that makes it necessary to perform several types of analysis, in order to give validity and final consistency to collected data.

There are two types of analysis involved in the coherence analysis performed to the results of the research: horizontal and vertical.

The historic records of prices and technical novelties that were applied to previous periods are examined through the horizontal analysis (carried out in the Territorial Branches). Furthermore, that the specifications of each item adjust to the parameters established in the investigation manuals is also verified. Equally, the previous and current absolute prices are analyzed. It is verified that the previous and current collected amounts are in relation or equivalence to the base unit indicated for each input (unit of measure); the registered variations are analyzed and the observations made by collectors and supervisors are validated, in order to justify the behavior of variations in prices and costs of each input, in accordance with the investigation performed with each source.

On the other hand, with the vertical analysis (carried out by DANE Central) prices and the minimum and maximum variations are appraised, an analysis of the behavior of prices and variations in the local and national scope is done, and the average prices and variations of the current and previous period are assessed.

Likewise, in the case of the CPI, the news context regarding market behavior expressed by the media, such as TV news, newspapers, magazines and online inquiries, is considered, and background information is used taking into account the objectives and scope offered by each research.

The results of the investigations of the Producer Price Index, the prices traded in the supply centers reported by the Price Information System for the Agricultural Sector (SIPSA for its acronym in Spanish), and the monthly results of the basket reported by the newspaper *Portfolio*, are used for the behavior of the food products group. For the total CPI, the inflation expectations survey is a point of reference.

Each of the most relevant results occurred in the month and at a national and city level, groups, subgroups, expenditure class and items per income level are reviewed and analyzed, after index calculation (output charts and index variations), for analysis of statistics purposes.

The monthly and historical variations, the year-to-date and twelve months, atypical variations, contributions and participations within the same period of the previous year, are compared in this framework.

After the index is calculated, all the variations that are analyzed reflect the average change in prices during a given period of time. These periods can be; monthly (monthly variation); year-to-date (year-to-date variation,) dating back years ago (historical variation) and between the reference month and the same month of the previous year (12-month variation). In addition to the variations, an expression that enables the procurement of the contribution (in percentage points) to the percentage variation of total prices, of each good or service, or any category of the CPI, is obtained, correcting the contribution with the price effect, of the previous period.

The previous calculations provided by the CPI research are used as a statistical tool for the thematic analysis on the evolution of prices and important price behaviors (ups and downs), at a local level, and are performed in order to determine conformity of the results. Finally, the results are compared with the context analysis that reflects the evolution of prices in the month, either locally or at a national level.

4.2. ANALYSIS OF QUALITY OF RESULTS

The indicators calculated at territorial level and others at central level are essential in the analysis of the quality of results. These indicators are followed-up each week prior to the calculation, delivery and publication of the index.

4.3. QUALITY INDICATORS AT TERRITORIAL LEVEL

4.3.1. Response rate Index or coverage by sources (ITRF for its acronym in Spanish). Represents the relation between the number of sources visited per month and the number of sources to be visited per month that was expected. The sources visited are considered to be those where the collector goes, even if the source does not provide the information, is undergoing a liquidation process, or in temporary closures. The objective of this indicator is to determine the responsiveness or “coverage” in terms of the sources visited or surveyed, versus the number of expected sources, or that were programmed or selected in the investigation. The calculation formula for this indicator is:

$$ITRF = \left(\frac{\text{Sources visited per week}}{\text{Expected sources per week}} \right) * 100$$

The reference level has 100¹¹ as its ideal value. Values below 92 should be supported with documents indicating the reason for this level and also the need for a corrective action should be considered. Values above 100% can occur when new sources are entered. Their source of information is the coverage report, the decentralized capture application - DANE local. It has a weekly and monthly frequency, and covers each city and aggregate.

4.3.2. Non-imputation or local estimation Index (INIL for its acronym in Spanish). This index represents the difference between the total number of records and the number of those records flagged with a technical novelty, brand or situation, that determine a process of imputation or some kind of estimation, and their relation to the total expected records by collection period. The objective of this indicator is to establish the level of imputation or estimation that an investigation is subject to, such as the difference between total expected records and the records marked for imputation. The calculation formula for this indicator is:

¹¹ This value is established in 100 in order to make relative comparisons that enable determination of the coverage behavior in terms of sources.

$$\text{INIL} = ((\text{RES} - \text{records for imputation or estimation}) / \text{RES}) * 100$$

The non-imputation or local estimation index (INIL for its acronym in Spanish), works with the following variables:

INIL= Local non-imputation index

RES = Expected records by collection period = number of items previously filled out for collection by period.

The reference level has 100 as its ideal value. Values below 92 should be supported with documents indicating the reason for the value and also the need for corrective action should be assessed. Its source of information is the Local DANE Decentralized application Report, with a weekly or monthly frequency, and covers each city or aggregate.

4.3.3. Local Quality Index 1 (IDCL1 for its acronym in Spanish). Quality of the chain processes Indicator that allows the production of investigations. The index is generated by the verification of the quality of work done at every stage of the chain: Collection and analysis, and calculated based on the detection of errors or omissions and penalizing them to the level of the smallest unit of information that is being measured (sources or records), but also considering the volume of well-done work at the same level of measurement. The form of calculation of this indicator is:

$$\text{IDCL1} = (\text{ICR} + \text{ICA}) / 2$$

Where: the local quality index 1 (IDCL1 for its acronym in Spanish) works with the following variables:

IDCL1= Local quality index one.

ICR= Collection quality index.

ICA= Analysis quality index.

The reference level has 100 as its ideal value. Values below 92 should be supported with documents indicating the reason for the value and also the need for corrective action should be assessed. Its source of information is the reports and the quality control forms of the process, with a weekly and monthly frequency, and covers each city or aggregate.

4.3.4. Local Quality Index 2 (IDCL2 for its acronym in Spanish). Is the quality indicator of the chain processes, which allows production of investigations. The index is generated by the quantification or sum of errors and omissions (nonconforming product treatment) at all stages in the chain: Collection and analysis, in respect to the total of records expected to be worked. The objective is to determine the level of quality of production processes of the research, such as the difference between the total expected records and the sum of nonconforming product treatments that occurred in the production process. The calculation formula of this indicator is:

$$\text{IDCL2} = ((\text{RES} - \text{TPNC}) / \text{RES}) * 100$$

The local quality index 2 (IDCL2 for its acronym in Spanish) works with the following variables:

IDCL2= Local quality Index two.

TPNC= Sum of omissions and errors in the processes of supervision and analysis' quality control.

RES= Number of expected records.

The reference level has 100 as its ideal value. Values below 92 should be supported with documents indicating the reason for the value and also the need for corrective action should be assessed. Its source of information is the reports and the quality control forms of the process, with a weekly and monthly frequency, and covers each city or aggregate.

4.4. QUALITY INDICATORS AT THE CENTRAL LEVEL

4.4.1. Response rate Index or coverage by records (ITRR for its acronym in Spanish). This index represents the relation between the number of records, or collected variables, with factual information per collection period, and the number of records or variables that were expected to be collected according to the previous fill out process per period. The objective of this indicator is to determine the degree of “effective” response or “coverage” in terms of records, compared to the number of records expected in the research. The calculation formula for this indicator is:

$$ITRR = (Factual\ records\ per\ decade / expected\ records\ per\ decade) * 100$$

$$ITRR = (RES - NT / RES) * 100$$

$$ITRR = (RE / RES) * 100$$

The response rate index or coverage by records (ITRR for its acronym in Spanish)) works with the following variables:

RE = Effective records = Information without technical novelties that cause absence of information (C.R).

RES = Expected records = number of items previously filled-out for collection.

NT= Technical novelties (C.R)

The reference level has 100 as its ideal value. Values below 92 should be supported with documents indicating the reason for the value and also the need for corrective action should be assessed. In some instances this indicator may be over 100%, because of new records and it covers every city or aggregate.

4.4.2. Non-imputation Index or central estimate (INIC for its acronym in Spanish). This index represents the differences between the total number of records and the number of records flagged as having a technical novelty, brand, or situation, which determines some kind of imputation process or estimation, and the relationship with the total of records expected per period; Technical novelty (PE) that is controlled at the central level, as an alternative of control for omissions and errors of local order. The objective of this indicator is to determine the level of imputation or estimation to which a research is subject to, such as the difference between the total of expected records and those checked for imputation in the central level about the information refined at local level. The calculation formula of this indicator is:

$$INIC = ((RES - Records\ for\ imputation\ or\ estimation) / RES) * 100$$

The non-imputation index or central estimate (INIC) works with the following variables:

INIC = Central non-imputation index.

RES= Expected records = number of items that were previously filled-out.

The reference level has 100 as its ideal value. Values below 92 should be supported with documents indicating the reason for the value and also, the need for a corrective action should be assessed. Its source of information is the databases of the research with a monthly frequency, and covers each city or aggregate.

4.4.3. Reliability Index (ICFA). This is the simple average of all indicators of the quality of processes in the chain that enables the production of the research, from those calculated at local level and Territorial Direction, to those calculated at the central level. The objective is to determine the quality level of the investigation's productive processes, such as the difference between the average of estimated indices and the determined reference level. The calculation formula of this indicator is:

$$ICFA = (ITRF + ITRR + INIL + INIC + IDCL1 + IDCL2)/6$$

ICFA= Reliability index.

ITRF= Sources Response Rate Index

ITRR = Records Response Rate Index

INL= Local non-imputation index

INIC = Central non-imputation index

IDCL 1 =Local quality index 1

IDCL 2 =Local quality index 2

The reference level has 100 as its ideal value. Values below 92 should be supported with documents indicating the reason for the value and also the need for corrective action should be assessed. Its source of information is consolidated in indices, with a monthly frequency, and covers each city or aggregate.

5 DISSEMINATION

5.1. DATA REPOSITORY MANAGEMENT

The consolidated database is kept in a dedicated server for indices at DANE Central. Queries and the necessary products for dissemination by means of press release-generating applications are created from this database. Currently these applications are in Visual FoxPro and Oracle.

The application for the central management of this database is administered by a user, whose profile is that of a production coordinator and, with the support of IT staff, the coherent generation of each of the information backup copies (back -up) is kept and revised.

5.2. DISSEMINATION PRODUCTS AND INSTRUMENTS

The dissemination products resulting from the research that are made public every month are:

5.2.1. Press release. This product presents the evolutions of certain aggregates of the month for the national total and cities by groups of goods and services, highlighting the variation of the basic expenditures that contributed the most to the variation of the total index. Its presentation is flexible and depends on the particular conditions of the moment.

5.2.2. Press Bulletin. In the Bulletin of Statistics, besides the information indicated for previous products, more disaggregated figures for the national total (by sub-groups, basic expenditures and classifications), by cities and levels of income are released.

5.2.3. Presentation. In the presentation the general results for the month, are shown; additionally the historical behavior of monthly, year-to-date and twelve-month variations are presented, as well as the information on the basic expenditures of more intense variations and greater contributions to the different types of variation. Finally, the overall results for each of the 24 cities are included here.

Likewise, the CPI output charts of the monthly survey results (Appendix A) are part of these instruments of dissemination.

Finally, the products that will be available to the community in general for consultation purposes, are:

- Guide for the use of the CPI.
- CPI methodology.

- Baskets for price-tracking purposes.
- Fixed -weights System.

This information can be found on DANE's website, which includes the link of *dynamic queries*, where users can obtain historical information on the index by applying selection filters. Also, the above-mentioned information is available in the entity's data banks.

Also accessible on the website are documents and information pertaining to the investigation, which are found in MS Excel® spreadsheets, or through the link of dynamic queries, which enable the user to obtain information, according to his/her needs.

In the data banks, DANE certifies the results of the research, and if required by our users the products that in terms of indices and percentage variations are produced monthly by the research, namely:

- Monthly, year-to-date, and annual variation.
- Monthly, year-to-date, and annual contribution.
- Index numbers December 2008 = 100 base.
- CPI spliced Series since 1954.
- National information (disaggregated by cities and by income levels).

RELATED DOCUMENTATION

In this section the documents related to thematic design including the CPI-08 methodology, consistency specifications and validation used as a guide to evaluate the information reported in the collection instruments (FUR and DMC) and the structure of output charts (Annex B), are referenced.

The statistical design methodology, which presents the methodological considerations followed in the imputation process of the CPI-08, as well as the methodologies for the design and selection of the sample sources for the index, the design of reliability and quality indicators that, as explained above, seek to ensure the production of the CPI-08 in optimal quality standards: the systems' design methodology that is required to produce the CPI-08: the manuals for filling out the forms or the collection manuals: and the analysis that guides the fieldwork done in the Territorial Branches, are also included.

Finally, the contingency plan for the analysis and mitigation of risks to which the research is exposed to, is presented in "Other manuals".

Thematic design

- Thematic design methodology
- Consistency specifications
- Validation specifications
- Output charts

Statistical design

- Statistical design methodology
- Review manual
- User manual

Other manuals

- Guide for the use of the CP

- Control formats instructions
- Contingency plan based on risk analysis
- Quality specifications handbook of the food products group
- Quality specifications handbook of the dwellings group
- Quality specifications handbook of the apparel group
- Quality specifications handbook of the health group
- Quality specifications handbook of the education group
- Quality specifications handbook of the entertainment, culture and recreation group
- Quality specifications handbook of the transportation group
- Quality specifications handbook of the Communications group
- Quality specifications handbook of the miscellaneous group
- Guide for the preparation of press releases on CPI

GLOSSARY

Annual variation: percentage change calculated between the reference period (t) and the same period of the previous year (year t).

Base period of the index: the period of time for which the index takes the value of 100.

Base unit: it is the unit used as a standard measure of the amounts associated with the observed price, when the equivalence of said price does not correspond to this unit. This price conversion is necessary, since prices can only be compared when they are expressed in the same unit of weight, volume or quantity (units).

Basic basket: representative set (or specific variety) of goods and services to which, the Bureau of Statistics, do compiling of prices.

Collection periodicity: frequency of collection.

Complementary source: it is a source that allows obtaining price information about a good for those cases when it is impossible to obtain information at the primary source.

Translation of Spanish definition: It is an alternative to make up for the absence of an item at a source in which a perfect substitution or a change of reference cannot be applied.

Context: an item of equal quality is searched for at another establishment, without any change in the characteristics of the item. The complementary source does not apply to any technical novelties. If the item or article does not appear again in the original establishment, information about a new article should be taken from the same source. If this is not possible, it must be taken from another source or it must be included in a different original source, where it is not quoted.

Conformity: fulfillment of a requirement.

Consumer Price Index (CPI): statistical indicator that allows establishing the average percentage variation in the prices of a set of final consumption goods and services demanded or used by households.

Consumption good: it is a good or service bought and used directly by the final user, which does not need any productive transformation.

CPI income: an operational classification where, by classifying the average household income, the target population is divided into three groups: Low income, which corresponds to 50 % of households in the classification, medium income which corresponds to 45 % of the households in the classification, and high income comprised by 5 % of the households with the highest average income of the classification.

Cumulative variation year to date/accumulated variation year to date: percentage change calculated between the period elapsed in the current year and December of the previous year.

Data Collection: surveys and all other methods of deriving information from different sources, including administrative sources.

Elementary aggregate: it is a relatively homogeneous set of goods both in their physical characteristics and in the behavior of price changes.

Elementary aggregate index: an elementary aggregate index is a price index for an elementary aggregate comprising only price data. (CPI Manual).

Immediate substitution: set to deal with the absence of a specification, or item, by replacing it by a “perfect substitute”, taking into account all the characteristics of quality, and using as an approximation to this concept the price, the quantity and the brand. The effect on the calculation system results in variations, in minimum ranges, that could be received by the consumer when making their consumption expenditure and moving in its structure of preferences.

Index number: an index number is “a quantity which shows by its variations the changes of a magnitude over time or space. Important features in the construction of an index number are its coverage, base period, weighting system and method of averaging observations.”

Item: An individual good or service in the sample of products selected for pricing.

Laspeyres price index: A price index defined as a fixed-weight, or fixedbasket, index that uses the basket of goods and services of the base period. The base period serves as both the weight reference period and the price reference period.

Nonconformities: non-compliance of a requirement.

Press bulletin: it is a official document that contains the main information about the research.

Press release (booklet): it is an official communication containing research data that will be released in a condensed manner. It can be defined as a summary of the press release and is disseminated to mass media.

Price reference period: the period that provides the prices to which the prices in other periods are compared.

Producer Price Index (PPI): this index measures the average variation of prices of a basket of goods (representative of domestic supply) in its first stage of marketing. The internal offer includes both goods produced domestically and imported goods.

Reference change (evolutionary goods): goods similar to or extensions of existing goods. When a good changes some of its main characteristics, the technical team must use this type of treatment to face this situation.

Tariff (fee): the prices of the goods or services which are set by a very restricted number of producers or distributors. Producers or distributors may influence consumers' habits by fixing tariffs in the market segment, according to the characteristics of consumers, for example, public services tariffs.

Unique Collection Form (FUR): the instrument designed for the collection of data according to the requirements of each research.

Waiting period: a collection period without information for any item or product, due to the absence of the brand or variety of the particular product whose price is being tracked. A waiting period is applied when the absence of information is temporary. Note: generally, only one waiting period can be applied.

Weight reference period: the period covered by the expenditure statistics used to calculate the weights. Usually, the weight reference period is a year. Usually, the weight reference period is one year.

BIBLIOGRAPHY

(Translation of bibliographic titles and names is for reference purposes only)

Boskin, M. (1996). *Toward a more accurate measure of the cost of living: final report to the Senate Finance Committee*. Washington D. C., United States: Diane Pub co.

Comité técnico proyecto nuevo IPC. (1997 y 1998). Reuniones técnicas, observaciones y comentarios a documentos del Proyecto (*Technical Committee new CPI project. (1997 and 1998). Technical meetings, observations and comments to project documents*). Bogotá, Colombia.

National Administrative Department of Statistics, DANE. *Metodología IPC-60 ("CPI-60 methodology")*. Bogotá, Colombia. DANE

_____. CPI Work Group. (2002). *Manual de Recolección IPC-98 ("CPI-98 Collection Manual")*. Bogotá, Colombia: DANE.

_____. CPI Work Group. (2002). *Metodología IPC-98 ("CPI-98 Methodology")*. Bogotá, Colombia: DANE.

_____. CPI Work Group. (1996). *Memorias de Seminarios del IPC, coordinadores, supervisores y recolectores ("Memories of CPI Seminars, coordinators, supervisors and collectors")*. Bogotá, Colombia: DANE.

_____. (1997-1998) *Project Evaluation Reports, Statistics Canada Symposium. (1997-1998)*

_____. (1998). *Guía de uso del Nuevo IPC ("Guide of the New CPI")*. Bogotá, Colombia: DANE.

Ducharme, L.M. (1994) *Re-engineering the Canadian Consumer Price Index: Proceeding of Statistics Canada*. Ottawa, Canada: Symposium.

Freire, E. *Informes del proyecto Nuevo IPC – Primer y Segundo semestre de 1977 ("Reports of the New CPI Project - First and second semester of 1977")* Bogota: DANE-Universidad del Valle.

International Monetary Fund. (2006). *Manual del Índice de Precios al Consumidor Teoría y Práctica. ("Consumer Price Index Manual, Theory and Practice")* Washington: International Monetary Fund. INE, Spain. (1993).

ANNEXES

ANNEX A. OUTPUT FRAMES OF THE CONSUMER PRICE INDEX (CPI)

The following list includes the main output frames of the Consumer Price Index (CPI):

- Table 0.** Index and variation by cities, according to income levels and national total.
- Table 1.** Index of the previous and current months, monthly variation and accumulated year-to-date and last 12 months, by income levels, according to cities national total.
- Table 2.** Indices and variation by groups of goods and services, according to cities.
- Table 2a.** Indices and variation by groups of goods and services, according to cities Low-income group.
- Table 2b.** Indices and variation by groups of goods and services, according to cities. Middle-income group.
- Table 2c.** Indices and variation by groups of goods and services, according to cities. High-income group.
- Table 3.** Index of the previous and current months, monthly, year-to-date and last 12 months' variation, by income levels, according to groups of goods and services, national total and cities.
- Table 4.** Monthly, year-to-date and last 12 months' variation, by income levels, according to groups of goods and services, national total and cities.
- Table 5.** Index of the previous and current months, monthly, year-to-date and last 12 months' variation, by income levels, according to groups of goods and services, national total and cities.
- Table 6.** Index and monthly variation by income levels, according to groups, subgroups, expenditure classes, and national total basic expenditures and by cities.
- Table 6.2.** Index and year-to-date variation by income levels according to groups, subgroups, expenditure classes, and national total basic expenditures and by cities.
- Table 6.3.** Index and 12-month variation by income levels, according to groups, subgroups, expenditure classes, and national total basic expenditures and by cities.
- Table 7.** Current index, variation and monthly contribution by income levels, according to groups, subgroups, expenditure classes, and national basket basic expenditure.

- Table 7a.** Current index, variation and year-to-date contribution, by income levels, according to groups, sub-groups, expenditure classes, and national basket basic expenditure.
- Table 7b.** Current index, variation and 12 months contribution, by income levels, according to groups, sub-groups, expenditure classes, and national basket basic expenditures.
- Table 11.** Basic expenditure, contribution and monthly participation sorted from highest to lowest by national basket contribution.
- Table 11a.** Basic expenditure, contribution and year-to-date participation, basic expenditure, contribution and year-to-date participation, national basket.
- Table 11b.** Basic expenditure, contribution and 12-month participation, sorted from highest to lowest by national basket contribution.