BUSINESS INTELLIGENCE FOR ANALYSING ISTAT HUMAN RESOURCES DATA BASE

Alessandra Dentini, Iole Zeppieri[1]

Abstract. In the recent years, the use of administrative data plays a central role in Official Statistics for reducing statistical burden. The internal administrative data could represent a strong tool to increase efficiency of management system. Using Business Intelligence techniques, the Italian Institute of Statistics (Istat) Duty Travel database is analysed in order to realise strategic information for stipulating agreements with transport suppliers. This innovative approach in Public Administration allows exploiting data that have a strong benefit in order to allow substantial financial savings to be allocated to scientific research.

1. Introduction.

The Italian Institute of Statistics (Istat) provides the official information for understanding and deciding the complex reality of Italy. Istat carries out surveys (exhaustive and sample) and collects data from different official sources. In this paper, an administrative database for internal use is treated to show the power of the Business Intelligence for increasing the efficiency of the management system. It is not just a tool for understanding the "status quo" of Istat Human Resources (HR) but it is a set of statistical analysis to construct a strategy for improving the "idea" of modern Public Administration.

The Business Intelligence (BI) is a technology-driven process for analyzing data and delivering actionable information that helps executives, managers and workers make informed business decisions. As part of the BI process, organizations collect data from internal IT systems and external sources, prepare it for analysis, run queries against the data and create data visualizations, BI dashboards and reports to make the analytics results available to business users for operational decision-making and strategic planning.

---

[1] The paper is the results of the common work of the authors. In particular, A. Dentini has written Sections 2, 4 and 5; I. Zeppieri has written Sections 1 and 3.

The duty travel information selected from administrative Istat database are made up of a several variables which include employee identification data (registration number, profile, staffing plan), duty travel administrative data (number of assignments, duration, destination, object, and institution visited, cost center) and finally, the detail of the cost items for single duty travel expenditure (type of transport, board and lodging). The data analysis covers the years 2009-2022 (thirteen years of time series!) even if the last years have not presented many duty travels due to Covid pandemic.

The first phase towards the construction of the database involved the transposition of the raw matrix to obtain the single duty travel for each record (row). The statistical analyzes used have extracted from a very large matrix of unstructured administrative data the information necessary to design strategic company policies aimed at saving and investing towards new strategies for enhancing professional skills. In fact, the ability to reduce the costs of missions through the signing of agreements with service providers is a goal that Istat will pursue in the coming years.

The paper is one of the first example of Business Intelligence (BI) in Istat, in which data purely used for administrative matters are used to obtain useful information for the management of activities, economic savings policies, relations with other institutional organizations and much more. We are talking about a wealth of information to be exploited to increase the effectiveness of some internal and external processes of the Istat. The paper is structured as follows. In the second section, the approach of BI in the Public Administration is presented; in the third section, the database construction is focused; in the fourth section the results of BI techniques are presented, and, in the last section, the concluding remarks are highlighted.

2. Business Intelligence in the Public Administration.

In addition, in the light of the regulatory reforms, the Public Administration is today engaged in a renewal process that focuses on the internal efficiency of each individual administration, on greater transparency and on more accessible, flexible and timely services, to facilitate the relationship between the public and citizen.

These levers drive the Public Administration to equip itself with tools for decision support, with regard to Business Intelligence as a solution capable of combining planning and strategic control, as well as the analysis of processes and the detection of recurring phenomena, with all the implications relating to their territorial dimension. The Business Intelligence (BI) is a technology-driven process for analysing data and delivering actionable information that helps executives, managers and workers make informed business decisions. As part of the BI process,

institutions collect data from internal Information Technology systems (IT) and from external sources, prepare it for analysis, run queries against the data and create data visualizations, BI dashboards and reports to make the analytics results available to business users for operational decision-making and strategic planning (De Vivo, Polzonetti, Tapanelli, 2011).

In recent years, business intelligence has developed to include more processes and activities to enable performance improvement. Such processes include a statistical approach.

• Data Mining: Use databases, statistics, and machine learning to uncover trends in large datasets.

• Reporting: Sharing data analytics with stakeholders so they can draw conclusions and make decisions.

• Performance metrics and benchmarking: Compare current performance data with historical data to monitor performance against goals. Typically, this is done using custom dashboards.

• Descriptive analytics: Using preliminary data analytics to understand what happened.

• Query Execution: Querying data with specific questions, for which BI extracts answers from datasets.

• Statistical analysis: starting from the results of the descriptive analysis, further exploration of the data using statistics, for example in relation to how and why a certain trend has occurred.

• Data visualization: Transform data analysis into visual representations, such as graphs, charts and histograms, for easier data consumption.

• Visual analytics: Exploration of data through visual representations to communicate information on the fly and follow the flow of analysis.

• Data Preparation: Compiling various data sources, identifying their dimensions and measurements, and preparing them for data analysis.

The Corporate Information System (URBI) collects, organizes, processes, and manages the data necessary for running Istat. Such data can be produced directly within Istat in the various business processes or be acquired because of relationships with external parties.
Similarly, such data can be intended for internal consumption or be intended for external purposes. URBI consists of a computerized part called Business Information System (BIS) and a non-automated part. The components (both computerized and not) of a Corporate Information System (URBI) can be divided into two categories according to their purpose:

    1) components for the support of operational activity. This part of the URBI is responsible for archiving, managing and processing all the information for carrying out daily activities.

2) decision support components. This part of the URBI concerns management, production and archiving to support managers in strategic choices.

Internally, URBI presents a series of data that are redundant and non-coherent, consequently they are not supportive for decision-making choices.

In this case, the "data" is different from the "information". In fact, the terms "data" and "information" are sometimes used interchangeably, but they are not the same. In general, data is defined as individual facts (without meaning), while information is the organization and interpretation of those facts: information puts those facts into context. Data is unorganized, while information is structured or organized. Information is an uncountable noun, while data is a mass noun. Data is not typically useful on its own, but information is. By carrying out operations of aggregation and interpretation of the data present in the Human Resources Database, we transform them into information useful for analysis and for the choice of strategic decisions.

Figure 1 represents one of the main objectives of BI, i.e., the enhancement of information starting from unstructured data systems. This enhancement translates, through a work process, into a sharp reduction in costs in management terms. We want to extract information from the vast amount of operational data present in corporate IT systems using two techniques: data warehouse and data mining. This does not involve entering new data into the corporate IT system but is responsible for reorganizing, integrating, filtering, aggregating data already available to obtain strategic information (or summary data). In detail, in our case study, the information system (URBI) contains a lot of information in an unstructured and inconsistent way so that the information value is very low, and the cost reduction objective is far away. By selecting a series of variables so that indicators can be constructed, it is possible to develop many statistical analyzes and, therefore, to produce reports useful for identifying intervention policies on Istat's organizational management. The last step involves the exploitation of these policies in order to undertake strategic decisions with a high information content for the Institute with a view to reducing costs and investing the savings obtained in other research activities. Therefore, thanks to BI, as the value of information increases, the cost of processes decreases.

Figure 1 – *The same data but…with more value*
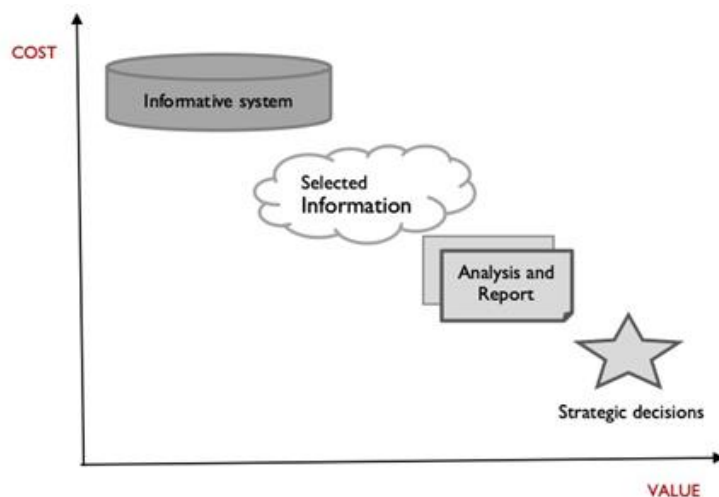
3. Data source and database construction

The Human Resources sector of Istat has been using URBI Smart as a management operating system since 2016. The application system manages various areas of human resources, including the legal and accounting management of employee duty travel and external staff who carry out activities for the Institute. In order to build a database that is statistically informative for the study of duty travel in Istat, it is necessary to review the work process divided into phases, starting from the rationalization of the information present in URBI in order to select the variables to be analysed for statistical purposes. The goal is to transform an administrative database, made up of multiple management variables, mainly used as a repository of administrative data, into a statistical database, capable of making the data classifiable and usable for statistical purposes. The duty travel information selected from URBI are made up of a several variables which include employee identification data (registration number, profile, staffing plan), duty travel administrative data (number of assignments, duration, destination, object and institution visited, cost centre) and finally, the detail of the cost items for single duty travel expenditure (type of transport, board and lodging). The data analysis covered the period 2009-2019 (eleven years) since this last year essentially represented the end of the duty travel due to the health emergency. The first phase towards the construction of the database involved the transposition of the raw matrix in order to obtain the single duty travel for each record (row) (Dentini, Mazziotta, Zeppieri, 2022).

Subsequently, the matrix underwent some changes summarized as follows:

• normalization of the cost items of duty travel for the years 2009 and 2015 when the management information system was different (SIGED was before URBI);

• normalization of the "Destination" variable in foreign countries and Italy, with a further subdivision by region, province and municipality;

• normalization of the "Destination" variable since the names of the destinations themselves were registered in a non-univocal way;

• normalization of the "Object of the duty travel" variable into four categories:
  o scientific meeting;
  o training;
  o internal meeting;
  o other institutional activities.

• classification by items of expenditure (travel, food and accommodation).

Table 1 – *Departures from Rome.*

Arrivals	Departures from Rome	…in percentage
Ancona	116	3,3
Bari	117	3,3
Bologna	474	13,5
Cagliari	109	3,1
Campobasso	63	1,8
Catanzaro	41	1,2
Firenze	415	11,9
Genova	104	3,0
Mestre	6	0,2
Milano	812	23,2
Napoli	473	13,5
Palermo	249	7,1
Perugia	98	2,8
Pescara	10	0,3
Potenza	6	0,2
Torino	159	4,5
Trieste	71	2,0
Cagliari	109	3,1
Venezia	179	5,1
Total	3.502	100,0

Our elaborations on Duty travel data base.

4. Results

In order to study the characteristics of duty travel, it is necessary to select from the database created a subset that focuses attention on the most important cases. In this regard, only the regional capitals are considered. Therefore, a double-entry table (not presented in the paper for reasons of size, 19x19) provides all duty travel to and from the cities hosting Istat regional offices. Just over twenty thousand routes in eleven years are considered. Of these, about eleven thousand routes insist on Rome, which, obviously, is the main headquarters of the National Institute of Statistics.

In table 1, the main destinations from Rome are presented; over time, 23.2% of the routes are towards Milan, 13.5% towards Naples and Bologna, 11.9% towards Florence. Palermo (7.1%), Venice- Mestre (5.3%) and Tourin (4.5%) present lower values than the cities already mentioned.

Table 2 – *Arrivals to Rome.*

Departures from…	Arrivals to Rome	…in percentage
Ancona	346	6,2
Bari	599	10,8
Bologna	330	5,9
Cagliari	319	5,7
Campobasso	156	2,8
Catanzaro	238	4,3
Firenze	477	8,6
Genova	275	4,9
Mestre	248	4,5
Milano	309	5,6
Napoli	487	8,8
Palermo	352	6,3
Perugia	252	4,5
Pescara	357	6,4
Potenza	263	4,7
Torino	249	4,5
Trieste	246	4,4
Venezia	57	1,0
Total	5.560	100,0

Our elaborations on Duty travel data base.

In Table 2, the main departure cities for Rome are presented; in eleven years, if we exclude Bari (almost 11%), Naples and Florence (almost 9%) which show outliers, all the other cities (including Milan) have a constant distribution of trips to Rome (about 5%). In fact, even the values of Venice and Mestre, if added together,

reach approximately this value. It seems interesting to study the situation of the route from Milan to Rome; in fact, the percentage (equal to 5.6%) is lower than other smaller cities that have smaller territorial offices. Probably because the Istat office in Milan (representing Lombardy) has such a high importance that it is more attractive than others. Moreover, for this reason, it organizes many events and has a higher number of incoming rather than outgoing duty travel.

Table 3 – *The cities most reached by Rome over time.*

| Departures from Rome	Bologna	Florence	Milan	Naples
2009	53	42	89	30
2010	48	41	56	20
2011	80	37	47	18
2012	55	24	52	62
2013	66	58	108	41
2014	40	30	71	59
2015	36	37	101	41
2016	23	32	57	75
2017	36	50	51	30
2018	19	25	66	48
2019	18	39	114	49
Total	474	415	812	473

Our elaborations on Duty travel data base.

Table 3 shows the major destinations reached by Istat workers whose place of work is Rome; these destinations are easily reachable by train, which is the most used compared to the plane, which is practically unused. The use of the train has also been encouraged by the increase in High-Speed Train (HST) over time. The frequency over the years towards these destinations appears to be constant.
There are peaks for Milan in 2013 and 2019; this phenomenon is due to two editions of Census of Companies that sees Milan as an operational hub.

The Table 4 shows that the most used vector is the train. With approximately 23,000 duty travel in Italy from 2009 to 2019, there is a constant trend between 64.2% and 78.2%, except for 2011 where the use of the train drops to 43, 5% and increases the use of plane (19.6%). This phenomenon can be connected to the Population Census was carried out for the last time in the traditional way. The greater use of the plane is also due to more destinations in the same duty travel.

Table 4 – *Trasportations in percentage*

Trasportations in percentage	Plaine	Train	Boath
2009	13,9	64,2	0,0
2010	12,6	69,5	0,0
2011	19,6	43,5	0,0
2012	8,1	64,8	0,1
2013	9,7	78,2	0,0
2014	13,2	74,8	0,2
2015	12,6	68,1	0,1
2016	10,6	76,3	0,1
2017	16,1	68,2	0,2
2018	6,6	62,1	0,1
2019	5,9	64,5	0,0

Our elaborations on Duty travel data base.

Figure 2 – *Of duty travel during and after Covid.*

Total Duty Travel
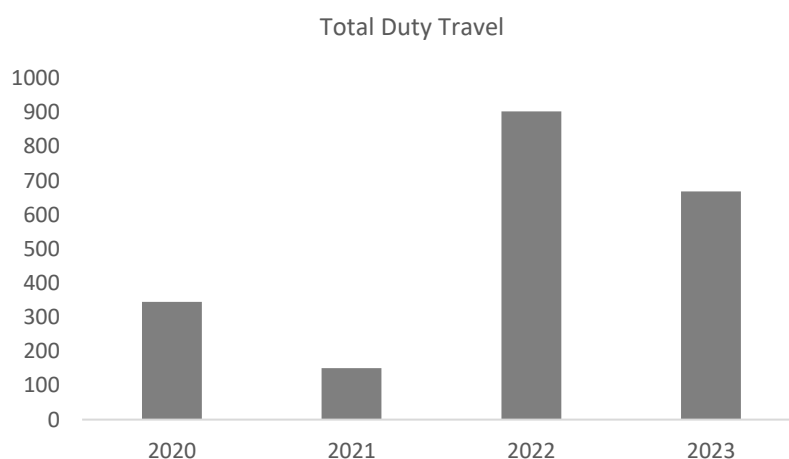
Our elaborations on Duty travel data base.

The SARS-CoV-2 pandemic emergency of 2020 prompted the Public Administrations to adopt gradual and timely managerial, organizational and technical choices for the protection of the health and safety of workers from this new risk, also limiting duty travel (Camisasca, Pietrantonio E., Magro, Arborea, Fabiani, Giordano, Pietrantonio M., Sparano, 2022). Before the pandemic, Istat researchers carried out around 3,000 duty travel in a year. The figure 2 represents duty travel

made in the pandemic and post-pandemic period. Duty travel made in 2020 mainly concern the pre-pandemic period (January and February 2020); a substantial decrease can be seen in 2021. Only at the end of 2021 does a weak recovery of duty travel begin, which increases more in 2022. In 2023, the data are encouraging, there is a clear recovery in travel with a projection for the end of the year of exceeding 2,000 duty travel.

5. Conclusions

The Italian and global Big Data and Analytics market is constantly evolving, with a high growth rate for both large companies and SMEs. In an economic phase dominated by an abundance of data and limited time to make decisions, it is important to exploit information to be competitive and generate business actions. Discover the advantages of Business Intelligence and Analytics to increase the productivity of your institution and simplify strategic choices.

Istat is starting a process of managerial growth in order to undertake activities for the coordination and use of strategies in order to provide technical-organizational support to the government structures of the Institute of Statistics.

This research experience must be understood as inserted within the context of profound innovation that statistics (not only official) are going through from the point of view of the use of administrative sources in order to represent complex realities with increasingly clear images that can assist the stakeholders in strategic choices.

The exploitation of the Istat duty travel database is a classic example of the use of administrative data for the development of BI techniques. In fact, BI refers to the ability to make better decisions, take informed action, and implement more efficient business processes. BI capabilities allow to: accept updated data from your organization; present data in easy-to-understand formats; provide data in a timely manner to make strategic decisions.

The objective of this contribution is to use internal Istat administrative sources for exploiting statistical information to rationalize internal processes and carry out economies of scale for improving the organizational structure of the Institute.

Besides, the final aim is the possibility of stipulating agreements with transport suppliers and with hotel structures for reducing the costs of duty travel and to invest the savings obtained in other research activities. The results of the statistical analyzes demonstrate that the number of duty travel has been constant over time, therefore allowing forecasts to be made on the future amount. Furthermore, the trend of the last year is strongly increasing therefore the prospect of returning to pre-pandemic levels has a very high probability. The analyzes agree in designing a scenario in

which the agreements to be stipulated can create considerable advantages in the organizational and financial management of Istat.

The first steps have been taken and have shown that statistical analyzes on databases for internal use can have a yet unknown value. The next steps guide the authors towards:

- interacting with Istat Mobility Manager;
- agreements with hotel facilities;
- analyzing the duty travel of external people and foreign destinations;
- implementing a data base available for consultation in real time;
- updating statistical reports to address decision strategies.

Official statistics, not only in Italy, are in a moment of great change. Administrative sources integrated with sample surveys are a new resource for providing quality statistics with less statistical burden for citizens and businesses.

Istat is acting in a new direction, exploiting internal administrative sources to increase process efficiency and financial savings.

Knowledge of BI statistical techniques will allow Istat to combine methodologies and unstructured data towards a new and modern working paradigm.

References

DE VIVO M.C., POLZONETTI P., TAPANELLI P. 2011. Open Data, Business Intelligence e Governance nella Pubblica Amministrazione, "Informatica e diritto", XXXVII Annata, Vol. XX, 2011, n. 1-2, pp. 239-262.

DENTINI A., MAZZIOTTA M., ZEPPIERI I. 2022. The destinations of research activity in Istat: statistical methods of business intelligence for enhancing territorial information, *Rivista Italiana di Economia Demografia e Statistica* - Vol. LXXVII-1 January-March 2023, pp 125-136

CAMISASCA M., PIETRANTONIO E., MAGRO R., ARBOREA A., FABIANI A., GIORDANO C., PIETRANTONIO M., SPARANO A. 2022. A pandemic health risk management model for the protection of workers: The Istat experience *Rivista Italiana di Economia Demografia e Statistica* - Vol. LXXVII-1 January-March 2023, pp 53-64

_____

Alessandra DENTINI, ISTAT, aldentin@istat.it
Iole ZEPPIERI, ISTAT, iozeppieri@istat.it