Business Intelligence.

The foundations of BI

Name: Kareen Ziadat - 20210004
Instructor: Dr. Rami Al-Ouran

CONTENTS

Background information about the department of statistics

Established in 1949, the department of statics carried out their first responsibility of Housing Units Census according to the Statistics Law No. 24. The current applied general law can be found on http://dosweb.dos.gov.jo/wp-content/uploads/2017/06/Statistical_Law_No_12_of2012-e.pdf. As time passed, the DOS expanded their statistical analysis to include a broader range of services such as expenditure and Consumer Price Indices (CPI), which were used to develop a Jordanian development plan. It later grew to include more industries, labor forces, population, and economic topics. The process of obtaining statistical data has been greatly facilitated as a result of technological advancements, in turn increasing the ease of decision making and, most importantly, saved time, money, and effort. One of the most significant changes occurred in the 2000s, when public statistical awareness increased, which increased the quality of data collected and increased trust. It was in 2010 when DOS was considered as a unified source of statistical data which allowed governments and private sectors to realize the importance of statistical data in analysis, decision making and research. As for the future plans, DOS is aiming to increase the quality and use of data collected in order to produce more informed decisions. All in all, the main goal of the department of statistics is to meet stakeholders' needs and allow sectors to make well-informed decisions by providing transparent, accountable, non-bias, and accurate data.

Type of business processes and supporting processes the department has or needs in the future.

Core processes: are the main processes that the organization does to achieve their ultimate goal. The main goal of the DOS is to collect and provide accurate and relevant data. Moreover, provided on their main website, the department of statistics has a list of the services they offer. These include carrying out comprehensive statistical processes on different sectors following a specific standard (standard used is dependent on the sector), providing statistical data to service applicants whether published, unpublished or raw), providing statistical publications issued by DOS (department of statistics), providing geographic statistical data maps that demonstrate statistics. They also train at the Jordanian Statistical Training Center (JSTC), help the service receiver conduct surveys by giving the manpower and expertise that help achieve their goal, and provide customer with same size. Along with more service options found in the link: http://dosweb.dos.gov.jo/DataBank/Services/Services2017_En.pdf.

Supporting processes: are the processes that help the core processes. In the case of DOS, these include: Defining and documenting the services provided. Maintaining their website by updating and securing it. Some services require approvals or requirements from external parties so obtaining theses approvals is essential. Human resources are extremely important as well to hire the correct people to do the job.

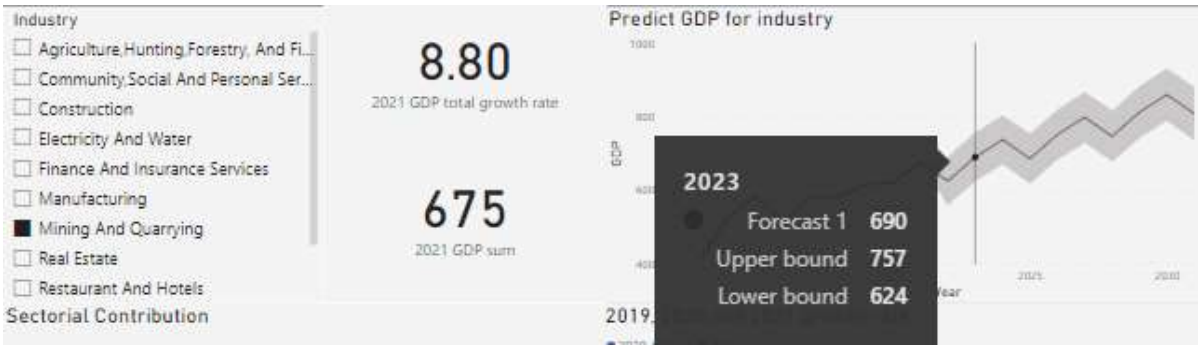
Management processes: needed to manage results, compliant. The DOS has made it clear that they focus on the data quality and therefore is considered a management process. They also make sure the standards of statistical data is met. A department for data mining is available to manage data provided.

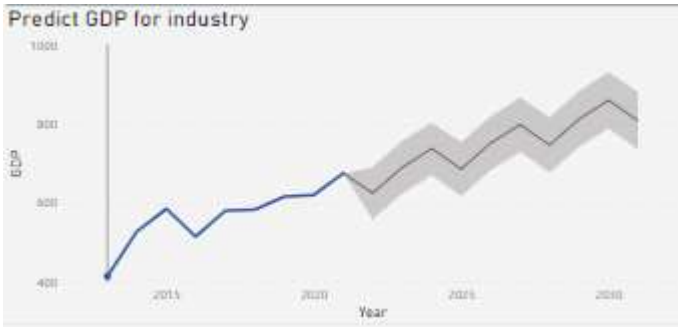
How proposed analysis help make decisions and main conclusions.

NATIONAL ACCOUNT DASHBOARD

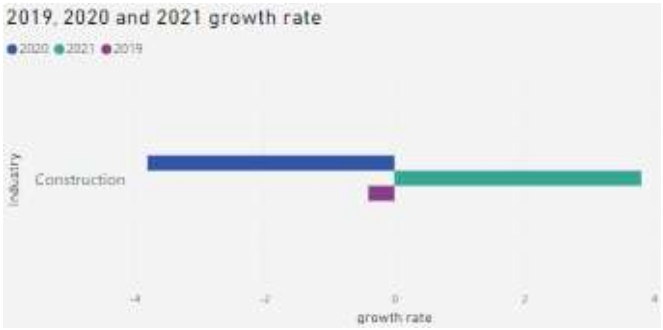
The first dashboard offers information about the GDP for each industry. GDP is an indicator for the economic performance of the industry. In my analysis, I used the GDP at constant price rather that at current price as it is adjusted to inflation which gives us a more accurate measure of the change in output (not increase due to inflation).

Organizations operating in the sector can choose their industry and obtain estimates of the GDP for the upcoming years. Investors will benefit from knowing whether or not the industry's economy is expanding and will be able to make their judgments and strategic plans accordingly. If the GDP is falling, the economy is deteriorating, and decreased profits are the result. Investors will create their yearly target objective, determine whether to boost pay, recruit more employees, and how much money to spend on their goods and services based on this information.

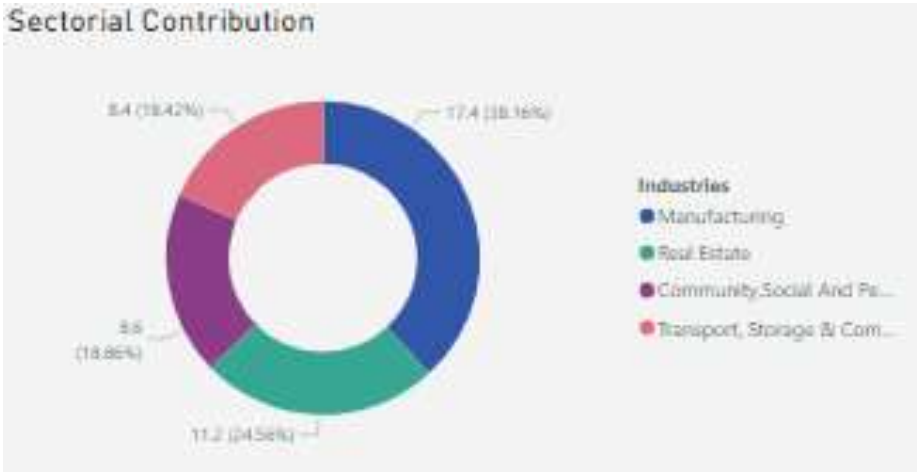




The most recent GDP and growth rate are also shown on the dashboard. The growth rate reflects how quickly the economy is expanding or contracting; ideally, we want the growth to be between 2 and 3 percent, which demonstrates that the economy is healthy and expanding. If further information about the growth rate is needed, a bar chart showing the GDP growth rate over the previous three years is used. This data is utilized to support the GDP estimate even more (increase or decrease confidence in the prediction). Investors can present all the provided information to stakeholders, evaluate the risk, and determine if the actions made based on the projections should be implemented.



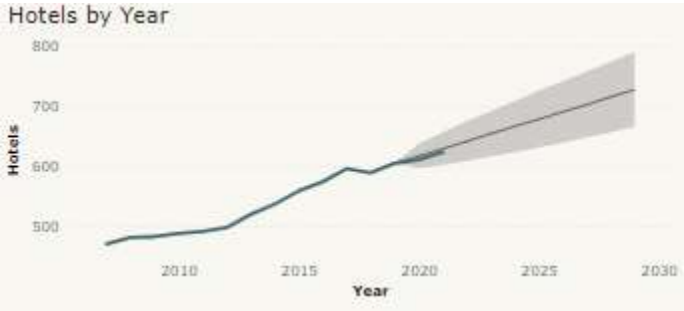
There is more information provided regarding the contribution of the sector to overall GDP. If the contribution is significant, the impact on the GDP will also be considerable, hence the sector has to be steady or developed if you want to keep or grow the GDP. The government may utilize this information to support the sectors that have a significant impact on GDP. These are the top industries that contribute to GDP:



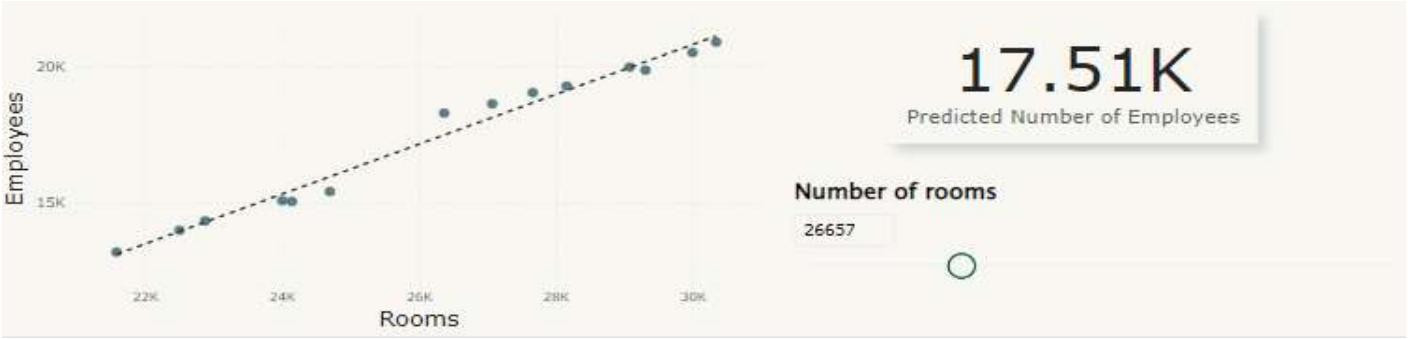
It can be concluded that it is wise to provide support (like increase money invested, education, etc.) to these industries: manufacturing, real estate, community and transport.

TOURISM

This dashboard concentrated on two key tourism-related aspects. The hotels and the tourist attractions and activities. The number of hotels may be projected for the following years, giving hotel owners a competitive edge by allowing them to plan for promotions, offers or expansions knowing that the number of hotels would rise.

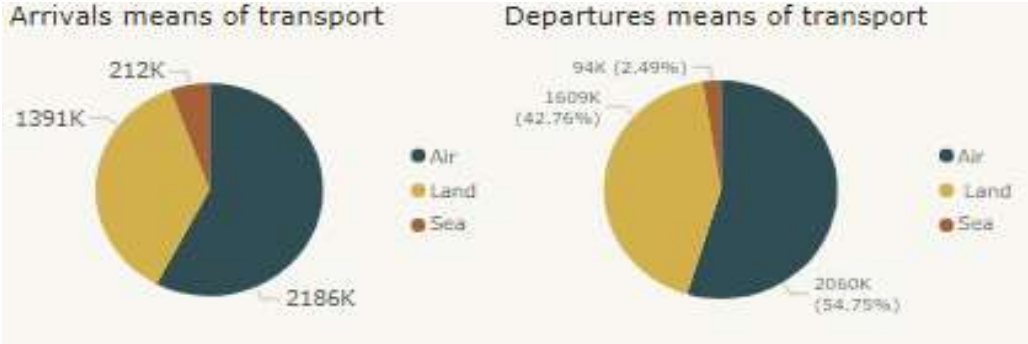


What-if analysis is used to predict the number of employees needed as the number of rooms of the hotels is changed (increased/decreased) this will allow us to anticipate hiring newer employees for the upcoming years or decrease the number of employees in case of a drop in the number of rooms due to a pandemic, loss of tourism, or other abnormal cases.

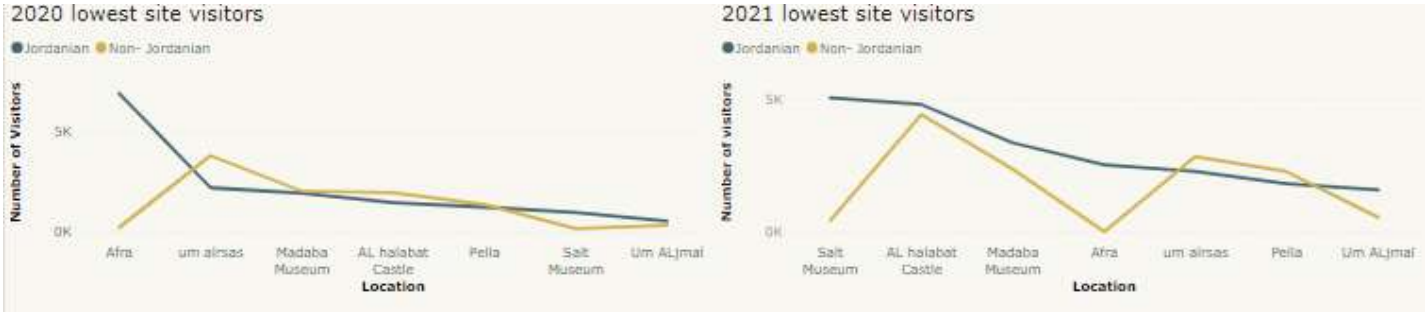


The number of predicted employees can also be used to inform job seekers about the demand for positions in the hotel industry.

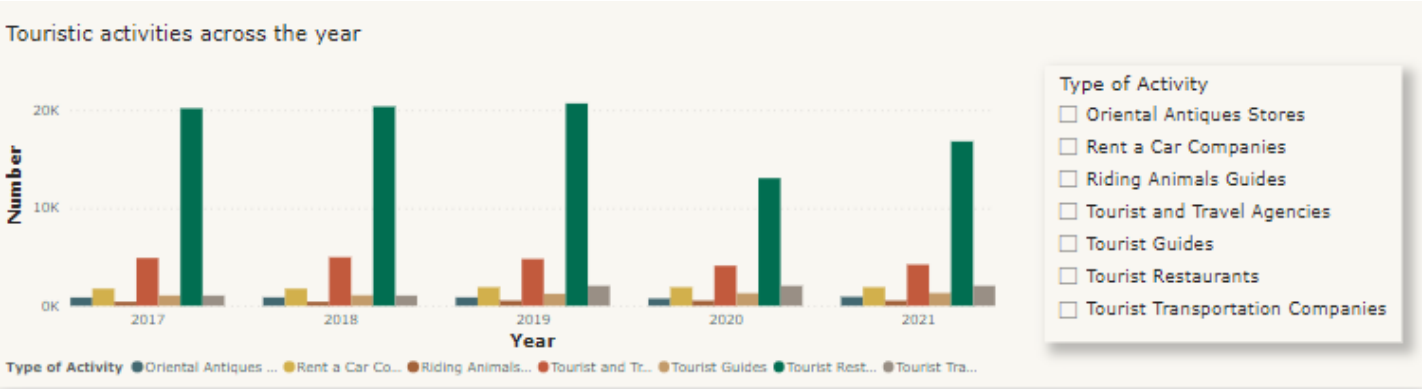
The dashboard indicates that flying is the most popular mode of transportation for departure and arrival. This indicates that airports require extra attention. On the other hand, we need to enhance travel by sea by investigating the reason behind the low number of arrivals and departures and putting a plan accordingly.



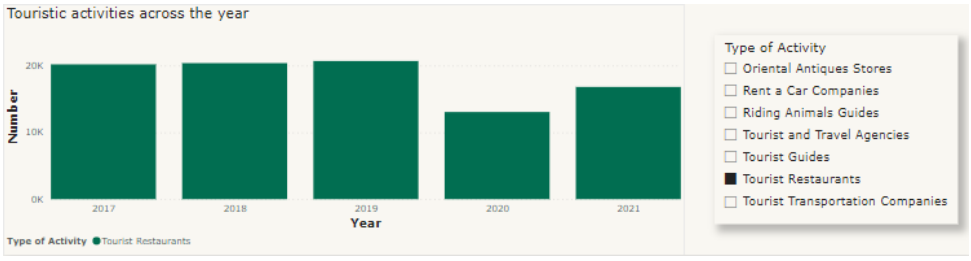
The second part of the dashboard shows the lowest visited sites by Jordanians and non-Jordanians during 2020 and 2021. Non-Jordanians tend to visit these locations less frequently, which may suggest that the reason these sites aren't visited is due to low exposure. Therefore, a plan should be put in place to increase exposure on these sites and to rehabilitate them to attract more tourist.



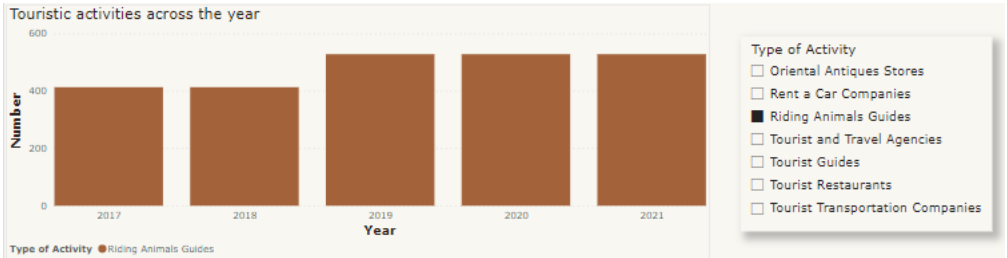
Lastly, a summary of the touristic activities by year (2017-2021) is shown. After selecting the type of activity, the user can see the number of the category and based on this information someone who is thinking of opening a business in this field can decide whether the market is saturated or if it's growing and it would be a good investment to initiate the business.



Ex: tourist restaurants is growing especially after the pandemic hit so it might be a smart decision to open a restaurant.



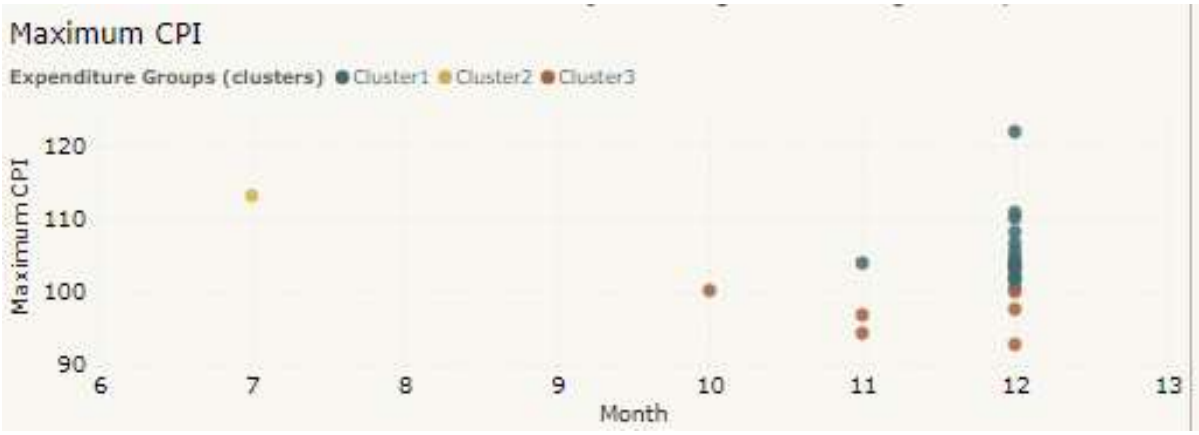
While the demand on riding animal guides might not be too high therefore it might not be a good idea to start a business providing this service.



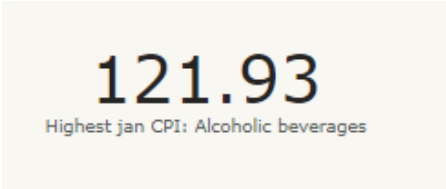
PRICES

In this dashboard I have focused on the consumer price index (CPI). The CPI is a measure of the change in the consumer prices compared to a baseline. When there is an increase in the CPI across the months, this indicates that the prices has increased over time. A high CPI can indicate higher inflation as sometime CPI can be used to measure inflation.

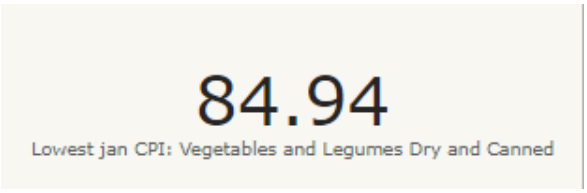
The dashboard showed that the maximum consumer price index usually happened during the month of December 2021, this can be because of the increase in the product prices due to the global economic crisis we faced.



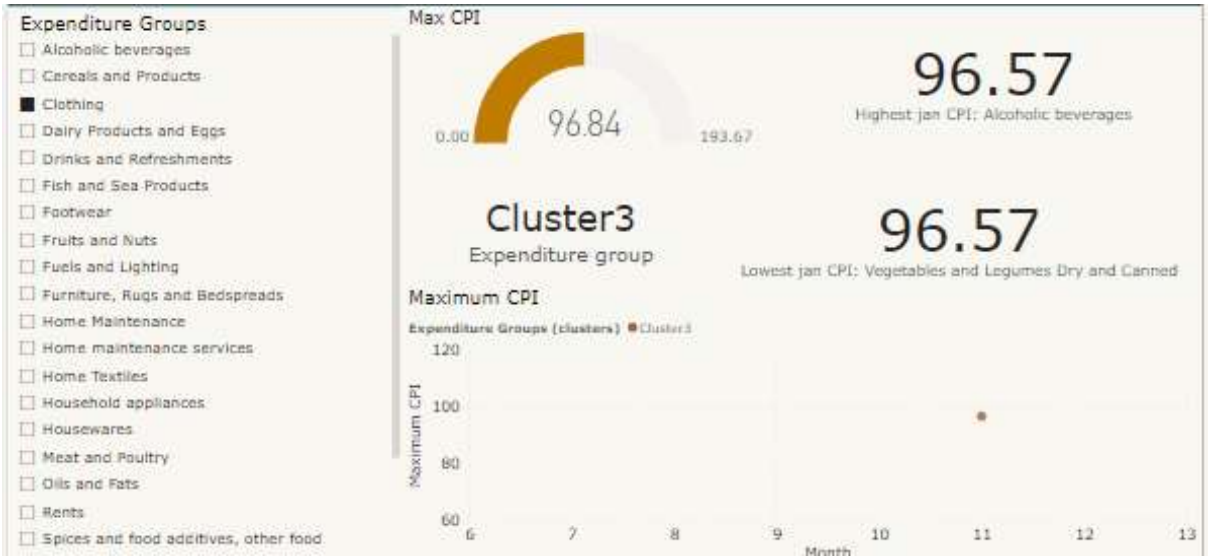
The maximum CPI across all expenditure groups was the alcoholic beverages. This means that the biggest price increase compared to the baseline (2018), was this group. This can be due to the fact that taxes have increased on this product as it is not a necessity.



On the other hand the lowest CPI was vegetables and legumes dry and canned. This indicates that the prices of these products haven't increased drastically compare to the baseline year which shows a positive indicator that these necessary product are still affordable.



The dashboard also provided clustering. The clustering was based on the maximum CPI for each expenditure group and during which month did this CPI occur. Three clusters were produced those with the highest CPI occurring in December or November and have a high value CPI. Products with the highest CPI occurring in October, November or December but with a lower CPI. And lastly which can also be considered as an outlier, the product that has the highest CPI in July. The business owner can select the expenditure group, see which cluster they are and make decision based on the cluster. For example, taking into consideration the clothing expenditure group. As a business owner who sells clothes, they will understand that the highest CPI is in November, meaning the inflation is the highest during that month which means that they need to increase the price of the clothes they are selling to avoid loss of profit. The price increase is dependent on the CPI value displayed by the gauge. At the same time since the product price increased, the demand decreases so a plan should be put in place to know the exact price increase and whether it is the right decision.



Types of data provided by the department of statistics.

Structured data: are data types that can be stored in a well-defined structure like columns and rows. Data types include numeric, strings and categorical data. The department of statistics mostly provides structured data in the form of spreadsheets (Excel) of statistical data concerning a wide variety of sectors. The DOS also uses databases that can be queried. The use of structured data makes it easier to search for data. Delete, modify or add data. And easier to store, secure and perform data mining on the data. Therefore it is extremely beneficial that the data provided by the DOS is structured.

Unstructured data: These are data types that cannot be stored in a structured manner and doesn't have a specific format. Including pictures, videos, documents, PowerPoints, reports and surveys. The DOS uses unstructured data in many ways. First many reports are provided to summarize important findings. These reports include graphs, strings of text, images and more. To add they use surveys to collect data to form the structured data. The use of unstructured data allows flexibility as we aren't bound by a certain format.it also allow scalability. On the other hand, it is hard to search for this type of data and difficult to secure them.

Semi-structured data: has some structure but cannot be represented in a relational schema. Data types include JSON files, XML, server logs, sensor outputs, emails and more. While I couldn't find semi-structured data offered by the department of statistics, the use of semi-structured data can be beneficial to analyze traffic in Amman for example. This will be done by gathering information for sensors and analyzing them. On the long run, this data can be used to diverge traffic.

Types of support available for business decision-making for the department of statistics department

Strategic decisions: This type of decision-making is characterized by the high importance and risk of the decision needed to be taken. Therefore, it is less frequent, costs a significant amount of money and requires thorough and careful analysis. For the department of statistics to ensure its continuity and achieve their objective, strategic decisions need to be taken. Examples of strategic decisions include selecting the correct sectors to gather data and statistics on to provide their customers with useful data. Another important decision is which organizations to partner with and what data they will provide. Moreover, from time to time the DOS should introduce newer services based on the demand.

Tactical decisions: This type of decision-making occurs more frequently with less risk and a lower impact on the business than the previous type (strategic). Nonetheless, it is still crucial and plays an important role in the stability of the organization and the achievement of the strategic goals. Examples of tactical decisions include, deciding the work schedules of the employees of the organization. To add, since the DOS need to comply with standards, over time these standards will change so decisions need to be taken to ensure compliance, data integrity and maintain trust with the customers. And, like every organization, the DOS will need to decide whether they need to hire more staff and who to hire

Operational decisions: This type of decision-making occurs even more frequently (almost daily) with a lower impact on the organization. Some of these decisions include offering customer service by answering inquiries about services provided and responding to suggestion or complaints. Whether to grant access of data to the company or individual that requested it (this may include approval from third party organizations). Risk management is also a frequent task that should be taken into consideration daily as risks arise all the time.

While the DOS needs to make countless of decisions, they mainly fall into three categories. Strategic decision-making concern long term and big picture decisions, tactical concerns mid-term and operational concern day-to-day decisions. The complexity of the decisions also differ between the categories where strategic produce more complex decisions, tactical produce less complex and operational decisions are the simplest. To add, the different levels of decisions are made by different level managers. Despite their differences, they all facilitate decision making and are interconnected. To further explain, operational decisions are decision made to achieve the strategic decision and support them therefore if strategic decisions are modified so are operational decisions. Operational decisions are based in tactical decisions which are made previously.

Analytics and data mining tools used in my analysis

The data analytics and mining tool that I mainly used were excel and power BI.

The first step of data mining was data cleaning. I used Excel to clean the data provided by the Department of statistics and extract the wanted data. The reason I used Excel to do so was because I am familiar with the tool and it is simple to use. First, moving and deleting columns/rows is simple. To add, in some case I transposed the data from columns to rows to allow me to carry the wanted and efficient data analysis within power BI.

In addition I used the maximum function to find the maximum value of a row as it is simpler carried out on excel rather than power BI. I also used the lookup function to find the month corresponding to the maximum value. These two new columns were used for analysis in Power BI. Moreover, looking at the simplified data in excel allowed me to have a clearer vision on the data provided and what the target of the dashboard will be.

I utilized the power BI desktop tools for data visualization. Power BI offers visualization tools that are effective, straightforward, and highly configurable. It enabled me to create dashboards that offered a simple, friendly, and interactive space that highlighted important data points. Data analysis was made easy for knowledgeable and non-informed users by data visualization. Some of the tools used for visualizations where use of pie chart to represent a part of whole (ex. Means of transport). Line charts were used to represent change over time. Bar chart were used for comparison of different categories. Donut charts are also used for the same purpose of pie charts. Another tool was the scatter graph which allowed me to show trends between two variables.

Outlier detection: for the analysis to be accurate, outliers should be omitted. In forecasting the number of hotels, the year 2020 and 2021 were omitted as they were an outlier due to the pandemic.

Clustering: this tool allows descriptive representation and groups data into clusters. It was used to group products with high CPI based on the year it had the highest CPI. This helped see clearly which month had the highest inflation for a group of products.

Prediction tool in power BI was used. This tool was used to predict the GDP of an industry to make informed decisions for the future. It was also used to predict the number of hotels for upcoming year.

Preformed what-if analysis: Linear regression was used to predict the number of employees based on the number of rooms specified by the user. What-if analysis can helps us understand the impact on the business if a variable is changed. In the case the number of rooms changed, the company will need to increase or decrease the number of employees (hire or lay off).

Jordanian businesses or companies which already used or may use the department of statistics data.

A list of partners are provided on the website of department of statistics on the following link http://dosweb.dos.gov.jo/ar/partners/. While some of these partners only provide the department of statistics with data, some benefit from the data provides. After calling the Department of Statistics customer service, it was confirmed that a variety of government ministries utilize this data like the ministry of finance, ministry of water and irrigation, National electricity company and ministry of transport. Universities, researchers and university students also benefit from the data provided. To add, many private companies use the data provided. One of these companies is Bloom Dead Sea Gift Enterprise.

How the data can help with their decision making or future investments

Bloom Dead Sea Gift Enterprise have obtained the exports and imports information and details about companies that are within the same sector. They obtained data about companies in their same sector to gain a competitive advantage and boost their sales on the long run. Information about exports allowed the company to understand where the product they're offering (Dead Sea products) is being exported to the most and whether Jordan exports enough. This allowed better targeting for countries especially when selling online.

In general private organizations can cluster data to target their product to a specific demographic to boost their sales.

Moreover, many news sources obtain their news based on data provided by the department of statistics. These news sources use the data and produce information that is helpful and important and can be used by the public. For example Roya published an article stating that the number of unemployment decreased.

Sources: https://en.royanews.tv/news/36173/2022-06-20.

https://en.royanews.tv/news/37227/2022-08-31.

https://www.jordannews.jo/Section-112/Economy/Unemployment-in-Jordan-stands-at-22-6-in-Q2-21272

Security or ethical issues with the collection and distribution of the data by the DOS.

The department of statistics uses a variety of data sources like third party companies, governmental institutes or other. The security of these data sources is extremely important as they are vulnerable to security breaches. Attackers can delete the data, modify the data or obtain the data and use it in unethical ways. Data loss will affect the processes of the department of statistics and their overall continuity. Moreover, if an unauthorized user was able to access the data, the confidentiality is affected and the privacy of the customer or the institute that provided the information is also threatened. Not to mention, attackers with bad intention can benefit from the provided data to cause harm to individuals or to companies. To add, if an attacker was able to modify data, the integrity of the data is affected and since organizations are basing their decisions on the provided data, data accuracy should be preserved to produce accurate decisions.

In case of a data breach, the DOS will have a huge financial loss and will face legal implications that are time consuming and damage the reputation of the DOS losing trust with customers. Customers' trust is extremely essential and being able to ensure the data of customers is safe will increase the customers' trust and encourage them to provide more data ensuring the continuity of the DOS.

It is important to note that despite the fact that the department of statistics complies with regulations, it is not enough to maintain the ethical use of data. A major security issue is the use of data incorrectly with the wrong intention of harming people. Since a huge amount of data is available to the public on the DOS website, any business or individual can access this data and utilize it to cause harm. A disclaimer is provided on the department of statistics website on the following link: http://dosweb.dos.gov.jo/disclaimer/ stating that the DOS isn't responsible for the misuse of data.

Another issue with data collection is the use of surveys which may not be reliable. In some cases questions may be misinterpreted or the survey was rushed and not taken seriously. This will affect the accuracy of the data presented.

Improving data collection and data distribution by the department

The use of data analytics software like integrate.io is recommended as it facilitates securing data. This software ensures compliance with GDPR or HIPAA standards and provides encryption for sensitive data ensuring the confidentiality. Another important aspect the DOS needs to take into consideration is restricting who has access to the data. This can be achieved using a method of authentication (1 factor, two-factor, etc.) and access control. The DOS also needs to improve data distribution by making all data private and only allowing access to data when the reason for data use is known and valid. And make sure to secure the transmission of data in order to maintain security, integrity and confidentiality.

Since the department of statistics relies on third parties for data collection, the integrity and confidentiality of the data provided by these organizations is extremely crucial so a process should be put in place to ensure that these organizations are collecting data accurately and securing them correctly.

References

dosweb.dos.gov.jo. 2022. No page title. [ONLINE] Available at: http://dosweb.dos.gov.jo/DATABANK/SERVICES/SERVICES2017_EN.PDF. [Accessed 18 August 2022].

Gramlich, M. (n.d.). What is structured, semi structured and unstructured data? › Michael Gramlich. [online] Available at: https://www.michael-gramlich.com/what-is-structured-semi-structured-and-unstructured-data/.

jorinfo.dos.gov.jo. (n.d.). PxWeb - Select database. [online] Available at: https://jorinfo.dos.gov.jo/Databank/pxweb/en/ [Accessed 16 Aug. 2022].

sisudata.com. (n.d.). Types of Decisions in Business Intelligence | Sisu Data. [online] Available at: https://sisudata.com/blog/types-of-decisions-in-business-intelligence.

smallbusiness.chron.com. (n.d.). Difference Between Strategy & Operational Decisions. [online] Available at: https://smallbusiness.chron.com/difference-between-strategy-operational-decisions-31075.html.

Smallcombe, M. (n.d.). Security and Business Intelligence: Why it Matters. [online] Integrate.io. Available at: https://www.integrate.io/blog/why-security-business-intelligence-matter/.