GL BAJAJ
Institute of Management & Research
Approved by A.I.C.T.E., Ministry of HRD, Govt. of India

Roll No…………

Plot No. 2, Knowledge Park-III, Greater Noida (U.P.) –201306

POST GRADUATE DIPLOMA IN MANAGEMENT (2019-21)
END TERM EXAMINATION (TERM -VI)

Subject Name:	**Data Visualization for Decision Making**	Time: **02.30 hrs**
Sub. Code:	**PGIT08**	Max Marks: **60**

Note:
All questions are compulsory. Section A carries 10 marks:5 questions of 2 marks each,Section B carries 30 marks having 3 questions (with internal choice question in each) of 10 marks each and Section C carries 20 marks one Case Study having 2 questions of 10 marks each.

SECTION - A

Attempt all questions. All questions are compulsory. **2×5 = 10 Marks**

Q. 1 (A): What are the Multi-Dimensional Data Plots? Mention the examples.

Q. 1 (B): How does Heat map show the data distribution better than scatterplot?

Q. 1 (C): How do we apply Gestalt Laws to data visualization? Give some examples.

Q. 1 (D): How is Python more suitable for Data Visualization? Explain with examples.

Q. 1 (E): What are Python Dictionaries? Describe their unique features?

SECTION - B

10 x 3 = 30 Marks

All questions are compulsory (Each question has an internal choice. Attempt any one (either A or B) from the internal choice)

Q. 2A: Write the Python code for a Creating Database Table.Create a new database DB1 and create a table with the name empmaster.

OR

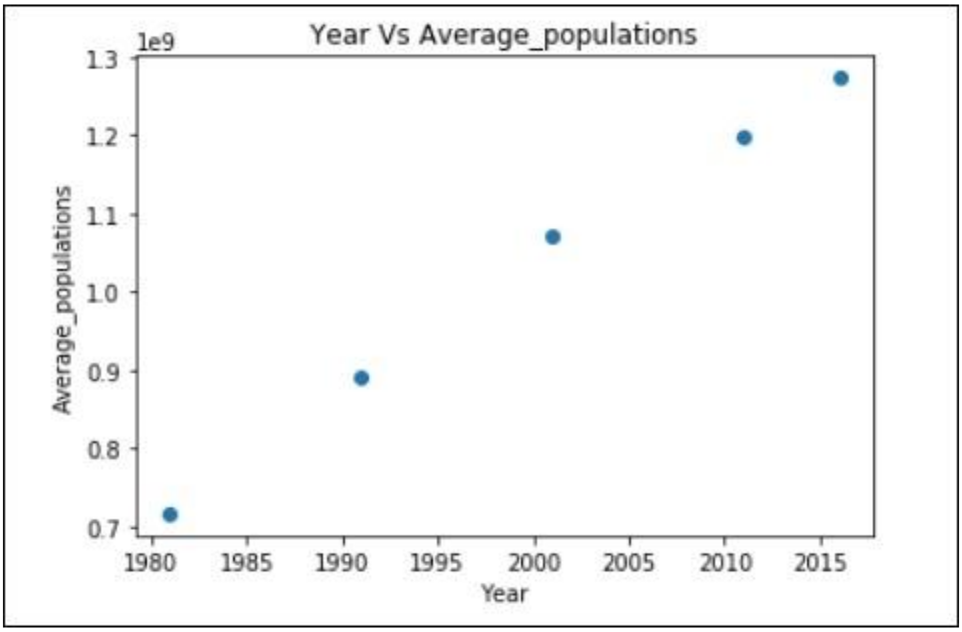
Q 2B: Write the Python code for Opening a file and appending some content to it.

Q. 3A: Write the Python Code for analyzing the Data File in CSV Format with multiple fields for Reviewing Skew of Attribute Distribution.

OR

Q. 3B: Write the Python Code for analyzing the Data File in CSV Format with multiple fields for Checking Dimensions of Data.

Q. 4A: Write the Python Code for creating the Pandas –SCATTER plot graph using matplotlib.pyplot directory and producing the output as given below:



OR

Question 4 B: For the Dataset given below, create the Pie Chart Data Visualizations.

Country or Area	Year(s)	Variant	Value
India	2019	Medium	1368737.513
India	2019	High	1378419.072
India	2019	Low	1359043.965
India	2019	Constant fertility	1373707.838
India	2019	Instant replacement	1366687.871
India	2019	Zero migration	1370868.782
India	2019	Constant mortality	1366282.778
India	2019	No change	1371221.64
India	2019	Momentum	1367400.614

SECTION - C

Read the case and answer the questions **10×02 = 20 Marks**

Q. 5: Case Study: Digitization and Visualization for Businesses

In many factories, digitization has primarily focused on efficiently operating the plant. Automation systems including PLC and HMI / SCADA systems control and provide a window into the automation equipment, and historical and traditional databases are used as a basis for production reporting. These systems focus largely on the current state (i.e., “What is happening?”) and on the historical view (i.e., “What already happened?”). This worked well for many years and drove significant improvements in industrial productivity as they became ubiquitous in most automated or semi-automated factories. The incremental benefit of these systems has largely been realized, and industry is searching for the next big digital transformation to drive further productivity.

Less than 1% of the data collected is being used productively today, for a variety of reasons. The data is often not structured well for extracting insights for the different groups involved. OEMs want to understand how their products are being used, to drive future product releases based on customer experiences. Operators and supervisors in the factory want to understand how different parts of the manufacturing process relate to each other to optimize the overall plant operation.

With the strong industry-wide focus on Digital Transformation and new products embracing the Internet of Things (IoT), we see the next major industrial transformation is already under way. Advanced sensors are being built into devices to collect data required to generate new insights, not just for purely operational purposes. Now, individual components of the assembly line can be automatically fine-tuned to work together precisely, performance in the field improves dramatically and economically through predictive maintenance that catches and remedies problems at the earliest stage. Designers can collaborate on upgrades based on actual usage However, the massive amount of real-time data that IoT devices yield can be overwhelming—a single jet engine with 5,000 sensors generates up to 10 GB of data per second. This creates new problems: how can manufacturers analyze data rapidly enough to use it, and how can people without data analysis skills understand and make use of it? Overall, the majority of manufacturers only use between 1 and 5% of their data.

Data visualization of the industrial IoT data and analytic results is the answer. It creates easy-to-understand and actionable graphic displays that can be used and shared throughout a company, supply chain, distribution network and to customers. In the past, every aspect of business was handicapped by lack of real-time, actionable data about how things—from assembly lines to tractors in the field—were actually working (or weren't). That data deficit forced business leaders to react only after a problem occurred. They had to rely on limited gauges read by workers who then manually recorded the data, which was only acted on later by supervisors. In complex assembly lines and manufacturing operations, this meant that every issue had the potential to cascade down the line before it was discovered. Even beyond production, because product designers had no idea how products were actually used after leaving the factory, they were forced to guess about wanted or needed features that would have the best impact in future versions. Likewise, maintenance costs were inflated as maintenance was done on a fixed schedule vs when failure warning symptoms were observed, and customers were still surprised by unplanned downtime as failures happened between maintenance cycles.

Using newer IoT techniques, customers can now assemble the data needed for a more condition based maintenance approach, where maintenance work is scheduled as soon as the pre-failure conditions are observed. Moving to this type of predictive maintenance strategy can dramatically reduce unplanned downtime, harnessing real-time data from built-in sensors to detect problems at their earliest stages and automatically trigger repairs, often before quality or productivity is affected. This increases reliability and customer satisfaction while reducing costs and catastrophic failures.

"Digital twins" of processes in the field let designers observe remotely how products are operating, facilitating targeted upgrades, while remote operators can use the twins to fine-tune operations. Some products can be improved on the fly with software upgrades. Manufacturers can even choose to share the real-time data with suppliers, supply chain, and customers, to optimize every aspect of the production and use cycle.

Too much of a good thing can be overwhelming. As more and more devices are equipped with sensors, the amount of data they generate can overwhelm your capacity to analyze and use it. IDC estimates 41.6 billion connected IoT devices by 2025 generating 79.4 zettabytes (ZB) of data. Others predict that the number of sensors may be increase to 1 trillion by 2030. Overall, we only use about 5% of the total data.

Much of this data isn't structured, is cluttered and disorganized, or is siloed in different departments, so people who all need it can't share access. Functional leaders lack complete access to data they need for critical decisions. They often must look to multiple sources that don't always communicate with each other. Also, not everyone in your company who needs this data to do their job more efficiently and/or make better decisions has analytical skills to make sense of it. Perhaps most importantly, if data isn't analyzed almost immediately, latency reduces its value—it becomes just another source of historical data.

Data visualization's effectiveness is partially due to how our brains work. Research shows visualizations are much easier and quicker for the brain to understand than data on a spreadsheet. It's also much easier to see relationships with a visualization than is possible with volumes of data in a spreadsheet, allowing correlations among various data sets that may affect each other, and uncovering solutions to one problem that may solve another. Similarly, teams from various departments and functions can break down organizational silos and easily discuss data at the same time, creating synergies and speeding decision making. It empowers individuals and departments that don't have data-mining and statistical analysts to act on data without assistance. Automated data visualization is a critical tool to eliminate the latency problem: you can act on the data almost immediately, improving operations and insight. Visualization also lets you use this real-time data to help identify emerging problems and opportunities quickly to make predictions and decisions in real time.

Question
Q5(A): Discuss the major challenge being faced by the modern businesses in the data handling?
Q5(B): How should organizations plan for their data management and visualization strategy? What are the potential benefits which may be reaped with data visualization?

Mapping of Questions with Course Learning Outcome

Question Number	**COs**	**Marks Allocated**
Q. 1:	**CO1**	**10 marks**
Q. 2:	**CO2**	**10 marks**
Q. 3:	**CO3**	**10 marks**
Q. 4:	**CO4**	**10 marks**
Q. 5:	**CO4, CO1**	**20 marks**

Note: Font: Times New Roman, Font size: 12.