DS 203 Assignment 4

Rohan Rajesh Kalbag
Roll no : 20d170033

Question 1

Python data types only contain information about whether the data contained in this is numerical/non-numerical and the algebraic operations that can be performed on it. Float can hold floating-point decimal numbers, and Int can hold integer numbers, and Object can hold non-numeric/alphanumeric data such as text.

Meanwhile, the statistical data types that we have seen in the lectures such as nominal/ordinal/continuous contain the relevant statistical information about the data and the non-algebraic operations that can be performed on it.

For example, Nominal/Categorical data contains data about a particular property of the entry such as name, address, gender which is of statistical importance. But object data type just contains information on whether the data contained inside is alphabetical/alphanumeric. Another example is ordinal data type which can be used to sort/rate the entry on urgency (for example in customer feedback data) but the corresponding data type of Python int just carries information that the contained data is numeric and nothing more.

Question 2

a. Descriptive
b. Predictive
c. Predictive
d. Exploratory
e. Exploratory
f. Exploratory
g. Exploratory

Question 3

a) Neonatal mortality means the death of an infant within the first 28 days of life, to identify some inferences on the difference between the city and rural scenario we do the following
    1) Exploratory: We obtain the data of the hospitals in the rural areas as well as the urban areas. Identify certain parameters such as the economic background of the parents, literacy of the parents, sex, weight, maternal age, pregnancy duration of the deceased infant.
    2) Descriptive: Identify parameters of the data that are common and more likely for neonatal deaths also try to identify correlations between certain parameters and eliminate redundant ones. For example weight and pregnancy duration of the infant may be positively correlated. Comparisons can be made between the urban and rural data.
    3) Predictive: Using the above inferences we can identify the most correlated parameters with neonatal mortality and conclude on the differences between the rural and urban scenarios.
    4) Prescriptive: Government policies can be made so as to address the most correlated parameters to reduce future neonatal mortalities, one example could be the distribution of Iron and Folic Acid nutritional tablets to pregnant mothers which would improve the health and immunity of foetus and reduce neonatal mortalities. Another example could be increasing the awareness of breastfeeding to improve the infant's immunity and overall health.
    Relevant Datasets/References:
    1)https://pubmed.ncbi.nlm.nih.gov/23734339/
    2)https://pubmed.ncbi.nlm.nih.gov/3251317/
    3)https://childmortality.org/
    4)https://data.unicef.org/resources/data_explorer/unicef_f/?ag=UNICEF&df=GLOBAL_DATAFLOW&ver=1.0&dq=IND.CME_MRM0.&startPeriod=1970&endPeriod=2021

b) Bellwether stocks are those stocks that are a leader in a particular industry and lead the market trends.
To identify bellwether stocks for a particular sector we need to do the following

1) Exploratory: Obtain the dataset of stock data for the firms in a particular sector (say technology) and sort the data in descending order with net market cap. Identify stocks with the largest intraday positive change as stocks with the potential to grow. Identify the top ~10-15 leaders of the market and analyse their market cap as a function of time.
2) Descriptive: Study common trends (parameters) in the market cap as a function of time in the leaders, find the range of intraday change these firms accommodate. Identify whether certain parameters such as stock price, volume and PE ratio are positively/negatively correlated. Identify if some outliers are there and study the reason for their deviation
3) Predictive: After identifying the necessary parameters now we can model an ideal bellwether stock and simulate the market cap and intraday change over a period of time using the model and predict which stocks are probable bellwether stocks.
4) Prescriptive: This data can be then used to influence/promote investment decisions of investors.
Relevant Datasets/References:
1)https://finance.yahoo.com/screener/predefined/ms_technology
2)https://finance.yahoo.com/screener/predefined/ms_technology/heatmap
3)https://www.investopedia.com/terms/b/bellwether-stock.asp
4)https://www.investopedia.com/top-tech-stocks-4581295

c) To choose a particular road over the others we compare the tradeoff between a variety of parameters. We shall follow the following guidelines.
    1) Exploratory: We shall need the data of the possible road projects that have been proposed. We shall also need the data of certain successful road projects of the past. We should identify certain parameters which are related to the economic growth achieved as a consequence of the road project such as the proposed no of vehicles travelling per day, the value of the goods transported per day also other parameters such as the entire price of the project, revenue from road toll, environmental and socioethical impacts such as the felling of trees.
    2) Descriptive: Study the common parameters that are present in the previously successful and socially accepted road projects. Identify the correlation between certain parameters and remove redundant ones.
    3) Predictive: Model a likelihood function for a successful road project using the conclusions and inferences made in the previous two steps. Evaluate the value of this performance parameter for each of the proposed projects and take the project with the maximum value.
    4) Prescriptive: Identify any drawbacks that may have been missed by the data, such as problems faced by residents, effect on biodiversity, land quality for infrastructure by actually conducting a physical survey of the land before finalizing the project.
    Relevant Datasets/References
    1)https://www.nbmcw.com/article-report/infrastructure-construction/roads-and-pavements/clearances-required-under-environment-acts-for-highway-projects.html
    2)https://www.sciencedirect.com/science/article/pii/S2352146518304484
    3)https://www.researchgate.net/publication/319488656_Success_Factors_of_Highway_Construction_Projects_in_Egypt_AHP_Approach