

DECLARATION: I understand that this is an **individual** assessment and that collaboration is not permitted. I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at <http://www.tcd.ie/calendar>. I understand that by returning this declaration with my work, I am agreeing with the above statement.

1.a) Exploratory Data Visualisation means exploring the data by visualising it through different charts, graphs etc. The key idea behind this step is to identify if there are any important features/parameters in the data and in some cases, we might not even find any such important feature. Therefore, exploratory data visualisation generally starts with just basic hypothesis in order to find if we can identify any important feature from data. We do not have any specific questions with regards to data but just general question. We then create visualisations from the data and try to identify if we can identify any relationship in the data from these features. Therefore, we can conclude that exploratory data analysis/visualisation is the process to explore data to identify any hidden trends in the data or if there any outliers in the data to identify any noteworthy features in data.

Explanatory Data Visualisation is usually the follow up step after exploratory data analysis, Basically, once we have identified any important features in the data while exploring the data, we use this step to use different visualisations to highlight the impact of these features in the data. We try to highlight how these parameters have impact in overall trends that we have identified from the data. These visualisations provide the specific and actionable insights from the data.

1.b) Data visualisation plays very important role in the data analysis process. It is key process to identify the hidden trends and key insights from the data that plays a significant role in identifying actionable items from the data. Visualisations from the data makes it easier to identify the important features from the data which are often impossible to analyse from the numbers. It is the most efficient way to identify these relationships from the data. By identifying these correlations and trends in data using visualisations, data scientist can build models that can be used to make meaningful predictions that can impact the decision-making process of the organisation.

2. I have used Covid-19 data of Ireland to plot a choropleth highlighting the number of active Covid cases per 100K population in different counties of Ireland as of 21 December 2020 and 21 December 2021. This choropleth shows which areas witnessed significant increase in the number of covid cases. I used R and ggplot2 library to plot this data. I have done this as part of online course for learning R.

Number of cases per 100K on the 21 Dec 2020

Number of cases per 100K on the 21 Dec 2021

