A2.1

Visualization 1: Tracking the flow of Military equipment to Police Departments

1. **Visual Encoding Channels**
    - **Colour**
      The left half of the image that shows the value of different military equipment procured is encoded using colour with each colour representing amount spent on set of equipment like vehicle, aircraft etc. The right half of image again uses colour to highlight regions in which different states of US lies and value of equipment received by that state.
    - **Colour Gradient and Size**
      The right corner of image shows bubbles of different size from small to big representing lower to higher value of military equipment per person. Also colour gradient from light to dark blue colour again shows lower to higher value of military equipment per person.
    - **Thickness/Width of Path**
      Width of trail in left half used to highlight the money spent on equipment, and on each category of equipment with thicker path representing more value. Similarly, thickness of trail on right side of image shows value of military good given to a state with more thicker meaning more value.

2. **Idioms**
    - Since this chart shows cost of equipment flow from different categories to different states, the main idiom used here is a Sankey chart.
    - Also, the right corner is a type of Bubble chart where different size bubble is used to highlight value per person.

3. **Dataset Type and Data Type**
    - There is Categorical data type in dataset like Equipment Type (Vehicle, Aircraft, Other), Regions (West, Southwest, etc.).
    - There is Continuous Quantitative data type in dataset like value per person, value of military goods for different equipment category and for different states.
    - The dataset type used in this visualisation is a multi-dimensional table. Some of the attributes of tables are value of military equipment, category of equipment, region, name of state, value of equipment procured by state, etc.

4. **Tasks and Comment/Analysis from the Visualisation**
    - Some of the key tasks associated with the visualisation are associate, categorize, cluster, distinguish, and rank.
    - Since this visualisation uses a Sankey chart idiom, it shows the flow of value of military goods to different states. This idiom used here is very intuitive and appropriate for this dataset.
    - The encoding channels used in this visualisation like thickness to show the value of equipment, colour used to depict various regions in which state lies and bubble chart representing value per person are very apt for this visualisation.
    - The left part of visualisation depicts the three main categories of equipment like Vehicle, Aircraft, and Other and the total value of these equipment.
    - Then this total value in each category is further divided into value of individual equipment in that category like MRAP in Vehicle category. Furthermore, it shows the value of that equipment as the percentage of the overall value across categories.
    - The Right side of visualisation shows the overall value of equipment received by each state in the US. It also highlights the region in which that state lies.
    - Finally, the visualisation shows the value of equipment per person in each state. This is calculated by dividing the total value of equipment received by state by its total population. It shows the value of equipment that Police department is getting for each person in state.

- The states are ordered in descending order of value of equipment it received which makes it easier to analyse.
- Also, highlighting value per person is a good statistic to analyse the value that state incurs to protect its citizens.

Visualization 2: Overnight stays in (US) National Parks

1. **Visual Encoding Channels**
   - **Colour**
   The visualisation uses 4 colour categories (Brown, Yellow, Blue and Green) for different lodging categories used by the visitors. Also, they have also used Colour encoding to represent temperature with white colour representing Cold Temperature (below 10°C) and red colour showing Warm Temperature (above 21°C).
   - **Thickness/Width**
   The visualisation uses width of radar to represent the number of nights the visitors spent in the National Parks.
   - **Position**
   Here positional encoding is used to indicate the month and the number of people that visit National Park.

2. **Idioms**
   - The visualisation uses radar layout idiom where the axes on concentric circles of the radar represents the number of visits and axes on the outermost circle shows the month of the year.

3. **Dataset Type and Data Type**
   - The dataset type used in this visualisation is a multi-dimensional table. Some of the attributes of tables are start and end date of visit, type of accommodation, national park visited, temperature etc.
   - There is Categorical data type in dataset like Accommodation Type (RV, Lodging, Tent, Backcountry), Season (Spring, Summer, Fall, Winter) and Month (January, February, etc.).
   - There is Continuous Quantitative data type in dataset like total nights spent and temperature.

4. **Tasks and Comment/Analysis from visualisation**
   - Some of the key tasks associated with the visualisation are associate, categorize, cluster, distinguish, identify, and locate.
   - The main analysis that can be done from the visualisation is they type of accommodation tourist prefer, season and time of the year tourist prefer to visit different national parks in the US. This will help anyone planning to visit a national park to plan accordingly.
   - The visualisation shows the total nights spent by the visitors across all the national parks in the US throughout the year which can be used to plan length of visit.
   - It also highlights the type of accommodation preferred by the visitors in different national parks.
   - It also shows the temperature and season along with the data using which we can also analyse its impact on the influx of tourist in various national parks.
   - Use of radial layout is appropriate since it is cyclical data and thickness of trails also helps user to identify popular means of accommodation easily.
   - Also, dividing the layout in different seasons also aids visitor to plan their visit judiciously.
   - Highlighting temperature along with season and total nights spent is also very helpful for visitors to plan their visit.
   - This visualisation can also be used to analyse different trends like accommodation preferences, seasonal preferences across various months in different national parks.

Visualization 3: World Population Data

1. **Visual Encoding Channels**
   - **Position**
   Here positional encoding is used to on the bar charts and the choropleth map.
   - **Colour**
   The visualisation also uses colour i.e., Orange, Blue and Olive as encoding channel for World Population, World Internet Users, and World Mobile Subscriptions, respectively.
   - **Brightness**
   Brightness has also been used as an encoding channel in Choropleths with darker shade of colour representing higher values and lighter shade of colour representing lower value. This pattern is followed in all 3 choropleths in the visualisation.

2. **Idioms**
   - Spatialized Data Arrangement in form of Choropleth Map and Bar Graph are the main encoding channels used in this visualisation. There are line graphs used in World Internet Users and World Mobile Subscriptions visualisation. This line can could be result of smoothing technique used along with the Bar Chart.

3. **Dataset Type and Data Type**
   - The dataset type used in this visualisation is a multi-dimensional table. Some of the attributes of tables are Year, Country, Population, Number of Internet Users, Number of Mobile Subscriptions etc. Since this visualisation contains choropleths, therefore the other dataset type in that is present is Fields as it will contain the position data for country as well.
   - There is Continuous Quantitative data type in dataset like Population, Number of Internet Users, and Number of Mobile Subscribers.
   - There is Categorical data type in dataset like Year and Country.

4. **Tasks and Comment/Analysis from visualisation**
   - Some of the key tasks associated with the visualisation are categorize, cluster, compare, distinguish, and rank.
   - One can easily identify the number of people in each category and also the trends in the data just by looking at the visualisation.
   - The Choropleth used as medium to highlight the number of people in different categories across different countries is very useful to analyse the data quickly.
   - The visualisation also depicts top 10 countries in each category in horizontal bar graph which is again very useful to analyse the data and trends.
   - The line graph generated using the smoothing technique makes it easier to highlight the hidden trends in the data.
   - The pop-up that comes on hovering the data makes it easier for the user to understand the values represented in different idioms used in the data.
   - The horizontal bar chart used to represent the year-wise growth in each category and vertical bar chart used to highlight numbers in Top 10 countries in each category are very easy to analyse from the visualisation

A2.2

Dataset: COVID-19 HPSC County Statistics Historic Data of Ireland

URL: https://data.gov.ie/dataset/covid-19-hpsc-county-statistics-historic-data?package_type=dataset

1. **Main data and dataset type in the data**
   - The main dataset used is Table, which is clearly visible from the data. The other data type is Field as positional data is also present in the dataset.
   - Attributes present in the dataset are: OBJECTID, ORIGID, CountyName, PopulationCensus16, TimeStamp, IGEasting, IGNorthing, Lat, Long, UGI, ConfirmedCovidCases, PopulationProportionCovidCases, ConfirmedCovidDeaths, ConfirmedCovidRecovered, SHAPE_Length, SHAPE_Area
   - There is categorical data present in the dataset like CountyName
   - There is Continuous Quantitative data type in dataset like PopulationCensus16, TimeStamp, IGEasting, IGNorthing, Lat, Long, UGI, ConfirmedCovidCases, PopulationProportionCovidCases, ConfirmedCovidDeaths, ConfirmedCovidRecovered, SHAPE_Length, SHAPE_Area

2. **Main Visualisation task relevant for this dataset**
   - We can perform various visualisation tasks like Rank, Distinguish, Compare, etc. We can perform various analysis tasks like identifying the trend of growth of Covid cases and Covid related deaths in different counties in Ireland. This will help authorities to plan effectively and optimally for better Covid management related contingencies plans. Also, we can use this data to analyse the number of recovered patients in each county and plan for additional medical assistance if required in any county.

3. **Visualisations on the dataset**
   We can do exploratory visualisation as well as explanatory visualisation on this data set. Some of the key visualisations that we can perform are:
   - **Time Series Analysis graph**
     We can make a time series analysis graph of all the counties in Dublin for the number of confirmed cases over a period provided in the data and check which counties have fared in containing Covid cases over the period of time. This can be a exploratory data analysis to find the tipping points in the visualisation that can help us to identify what factors might have causes that tipping point in the visualisation.

   - **Choropleth of all Counties in Ireland for Confirmed Cases, Confirmed Deaths, and Recovered Cases**
     We can use this visualisation to understand factors like Caseload per 100K population, Deaths per 100K population and Recovered cases per 100K population in different counties of Ireland. This is also exploratory data analysis.

   - **Bar graph of Daily number of covid cases in any county for a given period along with 7 day moving average**
     This can be used as explanatory visualisation to show how 7 day moving average progresses over time to highlight the effectiveness of initiatives taken by the government to contain covid. The average can also help to analyse and plan for future steps that needs to be taken.