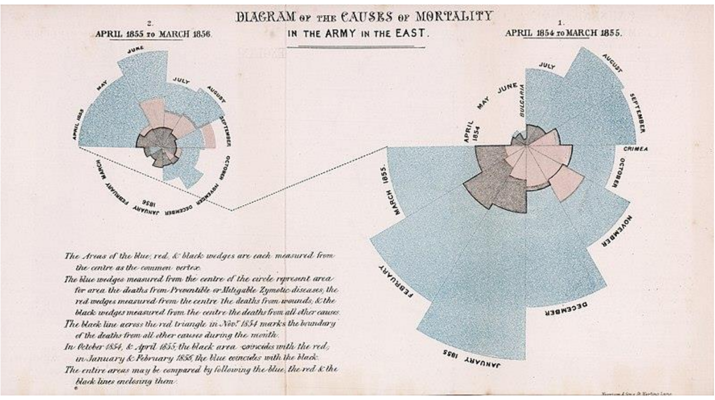
DECLARATION: I understand that this is an **individual** assessment and that collaboration is not permitted. I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at http://www.tcd.ie/calendar. I understand that by returning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

1 Part A: Florence Nightingale's "Diagram of the Causes of Mortality in the Army in the East"



This visualization is created by Florence Nightingale, which represents mortality rates in the British Army during Crimean War, specifically highlighting deaths from preventable diseases

From the chart we can infer that the dataset consists of categorical data (causes of death), quantitative data (number of deaths for each category) and temporal data (months from April 1854 to March 1856 in radial order)

The visualization employs polar area charts (also known as rose diagrams), a form of circular graph that utilizes wedges or sectors to represent data values. In this graph, the wedge areas are proportional to the statistics:

1. Wedge size: The most significant visualization encoding is through the area of each sector, representing the number of deaths due to three different causes including preventable diseases, wounds and other causes. The larger the area, the greater the number of deaths
2. Color encoding: Blue (preventable diseases), red (wounds) and black (other causes) color are used to encode different causes of death, making it easier to distinguish between the three categories.
3. Radial Position: Each wedge corresponds to a month from April 1854 to March 1856, allowing the viewer to understand how mortality changes over time. Months are ordered clockwise around the center, with two different charts for two years.

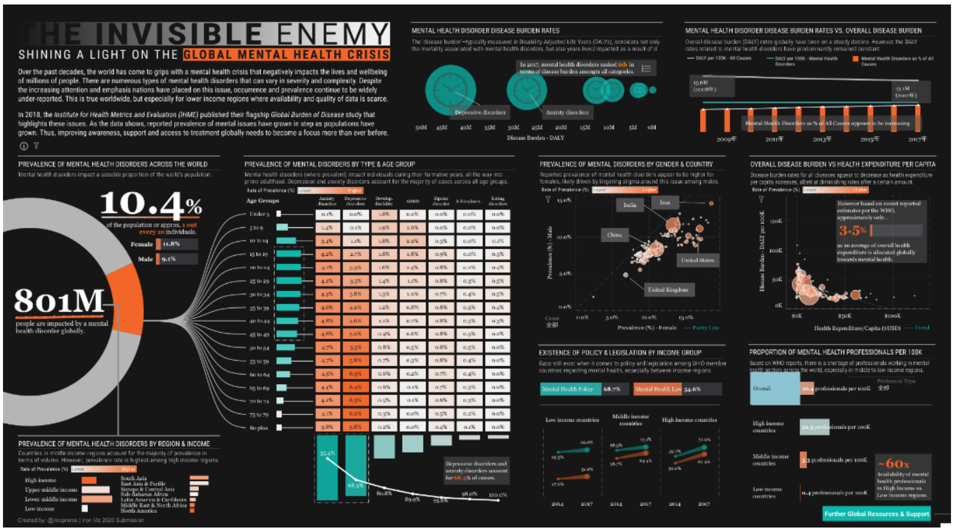
Several analytic tasks can be performed:

1. Identify dominant cause of death: Facilitates comparison between the three causes of death, making it super easy to spot out the most significant cause of mortality
2. Trend over time: Viewers can observe how deaths fluctuate in a year, hence easy to spot out the surge in death during winter months.
3. Compare between years: The two charts allow a comparison of mortality trends between 1854-1855 and 1855-1856.

Appropriateness of encodings:

1. The choice of using area to encode death counts is powerful for this kind of comparison, as the large discrepancies between causes are effectively highlighted. However, using area instead of length can sometimes be misleading, as humans are less accurate at estimating areas compared to linear scales.
2. The use of multiple polar charts is particularly well-suited to represent temporal data that repeats in cycles, like months of the year, because it highlights seasonal trends and allows for easy comparison across large time periods, oppose to standard bar chart which might not effective for showing periodic changes.
3. One potential drawback to use radial position and overlapping sectors is that data visualization for user would become harder to differentiate and reducing clarity. The graph could also hide the fine details about smaller wedges that close to the center of graph.

2 Part B The Invisible Enemy: Shining a Light on the Global Mental Health Crisis:



This visualization addresses the global mental health crisis, focusing on mental health disorder prevalence, disease burden and healthcare accessibility across various regions.

The dataset includes a mix of categorical, quantitative and geographical data. The categorical data includes the type of mental health disorders, age groups, income groups and regions. The quantitative data includes the prevalence rates, healthcare expenditure and the number of people affected globally. The geographical data identifies the prevalence of disorder across different countries.

There are several visual encoding techniques used across several charts:

1. The donut chart encodes the proportion of the global population affected by mental health disorders; The size of the circle also represents the number of people affected
2. Different Color is used for distinguish between gender groups and color intensity encodes the prevalence rate. Darker shades representing higher prevalence, extra bar chart making it easy to see anxiety and depression dominate across most age groups
3. The scatter plot visualizes gender-based prevalence in countries and overall disease burden versus health expenditure per capita (Position base). The size of the scatters corresponds to the population size while the color intensity represents the rate of prevalence
4. The bar graph breaks down the prevalence of mental health disorders by income group and region, which shows that low- and middle-income regions bear the highest burden of mental health disorders.

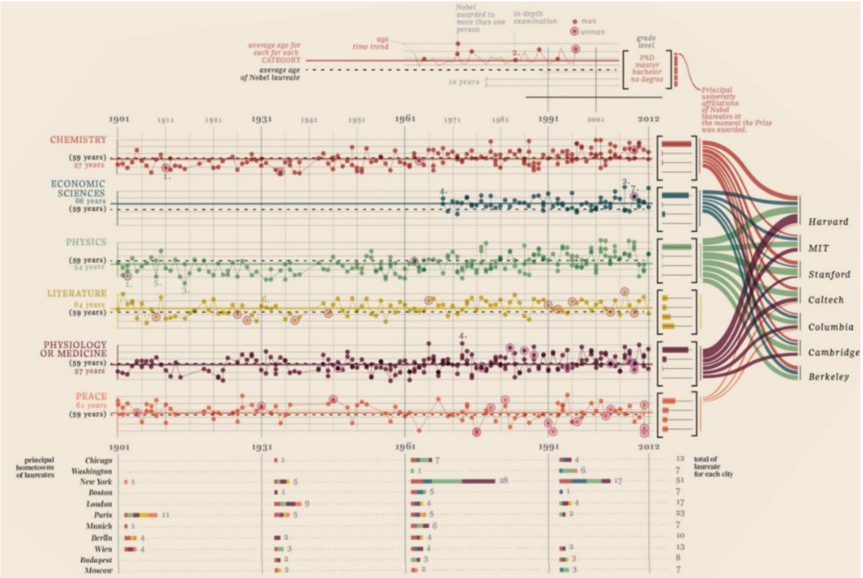
Several analytic tasks can be performed

1. The heatmap allows viewers to compare mental health disorder prevalence by both type and age group. This helps identify which disorders are most common for each age range.
2. The scatter plot comparing gender prevalence by country highlights where mental health disorders disproportionately affect specific genders
3. The scatter plot examining DALYs versus health expenditure allows viewers to explore how increased healthcare spending impacts mental health outcomes globally.
4. The bar chart breaking down prevalence by region and income groups offer insight into how economic factors affect mental health disorder rates.

Appropriateness of encodings:

1. The heatmap is effective in representing different category of data (age groups, disorder types, and prevalence) simultaneously, making patterns easily identifiable.
2. Scatter plots are appropriate for showing relationships between variables (gender vs. disorder prevalence, and health expenditure vs. disease burden), making it possible to detect clusters and outliers.
3. Circle sizes in scatter plots offer additional encoding of population, but can sometimes make precise comparison or data inspect challenging.

3 Part C Nobel Prize laureates over time



This visualization explores the story of Nobel prizes through years. Visualized for each laureate are prize category, year the prize was awarded, and age of the recipient at the time, as well as principal academic affiliations and hometown. Each dot represents a Nobel laureate, and each recipient is positioned according to the year the prize was awarded (x axis) and his or her age at the time of the award (y axis).

From the chart we can infer that the dataset consists of several dataset types:

1. Categorical data includes the prize categories (chemistry, economics, etc.) and academic affiliations (Harvard, Stanford, etc.)
2. Quantitative data includes the recipient's age at the time of receiving the award and the year the prize was awarded.
3. Geographical data is encoded through laureates' affiliations with institutions, providing insights into where the contribution is originated

The visualization uses a mix of dot plots and flow diagrams, utilizing color, position and connections to convey multiple data dimensions:

1. For Dot plots, each dot represents a Nobel laureate. X axis represents the year the prize was awarded (1901 - 2015), while the Y axis represents the recipient's age at the time of the award.
2. Color and text encode the prize category including chemistry, economics, physics, literature, physiology and peace, allowing easy differentiation between categories
3. Vertical spacing reveals age distribution within each category, so that readers can easily spot trends like younger versus older recipients in specific fields
4. The flow lines connect recipients to their principal academic affiliations. The line thickness likely corresponds to the number of laureates associated with the institution, revealing clusters of academic influence.
5. The bottom bar chart visualized the total number of laureates for each prize category, allowing viewers for quick comparison across different fields

Several analytic tasks can be performed

1. The position of each dot on the X-Y plane allows viewers to spot patterns in how age correlates with prize awards.
2. The lines connecting laureates to their academic affiliations allow viewers to see which institutions have produced the most laureates.
3. By positioning laureates along a timeline, viewers can examine the rise and fall in the number of awards in different fields.

Appropriateness of encodings:

1. The use of a dot plot for visualizing age vs. year is well-suited to this dataset, as it provides a clear timeline of laureates while also revealing age distributions.
2. The use of color to encode prize categories make viewers easy to distinguish between fields at a glance.
3. The flow lines can create some visual clutter, especially as multiple laureates are connected to a small number of institutions.
4. The choice of a bar chart is helpful for the dot plot, summarizing the total number of laureates per category and making it easy to compare between fields.

4 Part D Visualization Design

This dataset mainly focusses on Singapore's 24-hour Weather Forecast, API link: https://api-open.data.gov.sg/v2/real-time/api/twenty-four-hr-forecast

The dataset contains the following dataset types:

1. Categorical: date (the day for which the forecast is provided), forecast (weather conditions such as thundery showers), wind direction (primary wind direction such as SSW, ESE)
2. Quantitative: Temperature (minimum and maximum degree of Celsius expected for the day), Humidity (minimum and maximum humidity levels expressed as percentage), Wind Speed (Measured in km/h, indicating the range of wind speeds)

Possible visualization for this dataset

1. Comparison of Temperature Trends: Compare the daily high and low temperatures over time
2. Understanding Wind Speed and Direction Patterns: Show how wind speed and direction vary over the days
3. Correlation Between Humidity and Forecast Types: Explore if certain forecast types are correlated with specific humidity ranges
4. Multi-variable Analysis: Look at the interaction between temperature, humidity, wind speed and forecast types

Visualization Approach

1. Line Chart for Temperature Trends: A dual-axis line chart, with time on the x-axis and two y-axes (one for high and one for low temperatures), would allow users to track temperature changes over time.
   a) Encoding Channels: X-axis (time), Y-axis (temperature), color for differentiation between high and low temperature lines.
2. Wind Speed and Direction Visualization: A wind rose chart or radial bar chart can show both the magnitude and direction of wind. The length of each bar can represent the speed, and the orientation represents the wind direction.
   a) Encoding Channels: Bar length (wind speed), bar direction (wind direction).
3. Scatter Plot Matrix for Multi-variable Relationships: This matrix could allow users to explore the relationships between temperature, humidity, and wind speed. For instance, the x-axis could represent humidity, the y-axis temperature, and the size of the points could encode wind speed. Each point could be colored based on the forecast.
   a) Encoding Channels: Position (humidity, temperature), size (wind speed), color (forecast type).

	A	B	C	D	E	F	G	H	I
1	date	temperature_low	temperature_high	forecast	humidity_low	humidity_high	wind_speed_low	wind_speed_high	wind_direction
2	2024/10/11	25	33	Thundery Showers	60	95	10	15	SSW
3	2024/10/10	25	33	Thundery Showers	60	95	5	15	ESE
4	2024/10/9	25	33	Thundery Showers	60	95	5	15	SSE
5	2024/10/8	25	33	Thundery Showers	60	95	5	15	SSE
6	2024/10/7	25	33	Thundery Showers	60	95	10	20	SSE
7	2024/10/6	25	34	Thundery Showers	55	95	10	20	SSE
8	2024/10/5	25	33	Thundery Showers	60	95	10	20	SSW
9	2024/10/4	24	34	Thundery Showers	55	95	5	15	WSW
10	2024/10/3	25	34	Thundery Showers	55	95	5	15	VARIABLE
11	2024/10/2	25	34	Thundery Showers	55	95	5	15	VARIABLE
12	2024/10/1	25	33	Thundery Showers	60	95	5	15	SSW
13	2024/9/30	25	33	Thundery Showers	60	95	5	15	SSW
14	2024/9/29	25	33	Thundery Showers	60	95	5	15	S
15	2024/9/28	25	33	Thundery Showers	60	95	5	15	SSE
16	2024/9/27	25	34	Thundery Showers	55	95	5	15	SE
17	2024/9/26	25	34	Thundery Showers	55	95	5	15	SW
18	2024/9/25	25	33	Thundery Showers	60	95	5	10	VARIABLE
19	2024/9/24	24	34	Thundery Showers	55	95	5	10	VARIABLE
20	2024/9/23	24	34	Thundery Showers	55	95	5	15	VARIABLE
21	2024/9/22	25	34	Thundery Showers	60	95	5	15	SE
22	2024/9/21	25	34	Thundery Showers	55	95	5	15	SSE
23	2024/9/20	23	33	Thundery Showers	60	95	10	15	SSW
24	2024/9/19	26	33	Showers	60	95	10	20	S
25	2024/9/18	26	34	Partly Cloudy (Day)	55	95	15	25	SSW
26	2024/9/17	24	34	Light Rain	55	95	15	30	SW
27	2024/9/16	26	34	Partly Cloudy (Day)	60	95	15	30	SW
28	2024/9/15	24	33	Thundery Showers	60	95	15	30	SSW
29	2024/9/14	24	33	Thundery Showers	60	95	15	30	SSW
30	2024/9/13	24	33	Thundery Showers	60	95	15	30	SSW