This PDF file aims to link the target audience and problems that we identified in COMP0035 to provide the context of the project. In addition, it will explain and evaluate our choices on visualisations for this Dash APP based on the context.
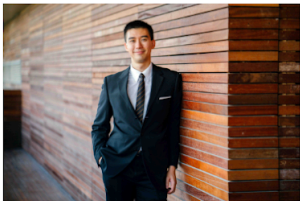
1. Background & Target Audience

As identified in previous courseworks, our app focuses on two types of target audience: employers and prospective employees (applicants). It is intuitively understandable that job applicants, especially female ones, want to find an appropriate job without discrimination against gender. Meanwhile, as we have already discussed in COMP0035, gender pay gap can ultimately have negative impacts on company performance as women counts represent half of the labor force and inequality against them can result in lower productivity. (Mollerop, 2019) Furthermore, the gender pay gap may also influence companies' attractiveness to high-skilled workers.

In this case, both user groups have demand for solving the gender pay gap situation. Therefore, we aim to design an APP that can provide data analytics for users to evaluate the pay gap in various aspects (Salary, Bonus, and Percentage of Highly Paid Female) against various dimensions (Region, Employer Size, and Industry). Hopefully this would help applicants to choose where to go and help employers to understand their gender pay gap position compared to others so that they can improve.

Personas were created for both types of target users to help emphasise the problem:

PERSONA FOR EMPLOYER (CHO) - JACKSON WANG

Challotte Wilson

Age: 27
Education: Msc Management
Work: Senoior Management consultant
Family: Newly married
Location: London, UK

Personality

Introvert — Extrovert
Analytical — Creative
Loyal — Fickle
Passive — Active

Ambitious   Focussed
Free-Spirited   Logical

Desired Companies

accenture
BAIN & COMPANY
KPMG
pwc

Bio

Challote entered management consulting through Accenture's female leadership program since 2018 and graduated with a first class degree in Msc Management from UCL. She is a professional management consultant who has the expertise that a company lacks. She has both theoretical and also practical experiences, so that she is able to help a business to focus on the development process and overall operations. Currently, she is looking to jump ship because she wants to move to her new married husband's city.

Motivation

Solving problems
Female-friendly
Personalization
High salary
Convenience working place
Clear career path

Preferred Channels

Linkedin
Glassdoor
Social Networks
Headhunters

"I am looking for employers who have prioritised gender equality in the workplace!"

Wants and Goals

- Get the same opportunities as men in the work place
- To build relationships with long-term clients
- To get more leads and grow her own business later
- To find a new female-friendly company in a safe way

Frustrations

- Personal information leaked during job change
- Have to face discriminations and unequal treatment because of stereotypical expectations, attitudes and behaviours towards women

2. Questions to Be Answered

Potential queries from our target users are listed below.

For job applicants:

- How serious is gender pay gap in general?
- What is the average percentage difference between male and female in mean of hourly pay and bonus pay of each region, employer size and industry?
- Which industry and company size are more female-friendly with higher percentage of highly paid female employees?
- Which industry / employer size / region has the biggest gender pay gap?

For recruiters:

- How serious is gender pay gap in general?

• How is my company's gender pay gap compared with others in the same dimension (industry, region, and employer size)?

3.   Visualisation Design

In order to answer questions identified above, we decide to create four types of visualisations which can present gender pay information from different aspects to users by selecting what they want to compare with selector and dropdown boxes. This section will explain and evaluate our choices on visualisations.

3.1 Histogram
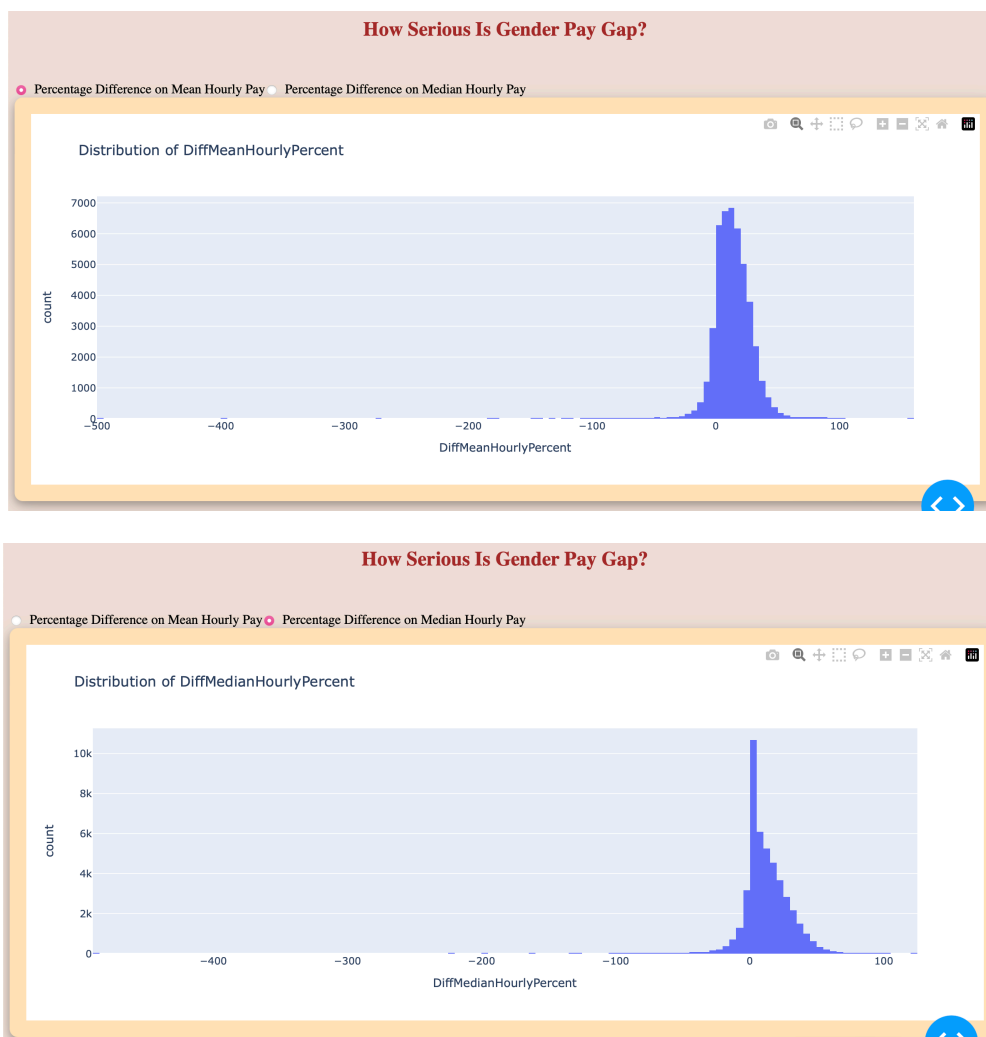
Two histograms were used to answer the question of **How serious is gender pay gap in general?**

Explain the design:

To address the above problem, we determined that a histogram would be the most effective graphic because histogram displays the distribution of continuous data. It places data into ranges of values known as bins and each bin represents the count or percentage of observations that fall within that bin, (Frost, 2021) so that users can easily understand the distribution. In this case, it enables users to view the distribution of gender pay or bonus gap in terms of mean percentage difference or median percentage difference. As the data chosen represents the percentage

difference on payment to male and female where larger numbers mean larger gap in favour of male, the more left-skewed the graph is, the more serious the gender pay gap against women is, and vice versa. Therefore, it is able to answer the question. Meanwhile, as the graph only aims to provide a general information, we have only included data represents hourly payment, which directly shows the payment situation.

Evaluate the design:

The histogram provides a clear and easy-to-understand representation of the data, allowing users to quickly identify patterns and trends in the pay differences. Users are able to compare the distribution of percentage difference in gender pay with the skewness, identifying the central tendency, including the mean, median and mode of percentage difference on hourly payment. This gives users a general information of the situation in the salary pay for different genders. However, only the mode is most obvious as the highest bar, we did not indicate mean and median on this distribution, which can be improved by adding vertical lines.

We added a functionality for users to switch the histogram in either the mean difference of the median difference on top. Users are able to compare the shape for the distribution of histograms as the median pay is often considered a more representative measure of central tendency, as it is less sensitive to outliers in the pay distribution.

However, histogram have contextual limitations that provides only a snapshot of the pay gap at a specific point in time, and may not reflect changes over time or differences in pay between different organisations. To improve this, we may need to combine data from multiple sources to create a more comprehensive view of the gender pay gap. This could include official data from government, private organisations. When presenting histograms of the gender pay gap, it is important to provide context and information about the data used, such as the time period covered, the size of the sample, and the method used to calculate the pay gap. This could be done by adding a section below the graph with a short paragraph summarising the pattern of histogram in words and figures.

3.2 Box and Whisker Plot

The Box and Whisker Plot is used to answer the question:

- **What is the average percentage difference between male and female in mean of hourly pay and bonus pay of each region, employer size and industry? (For applicants)**
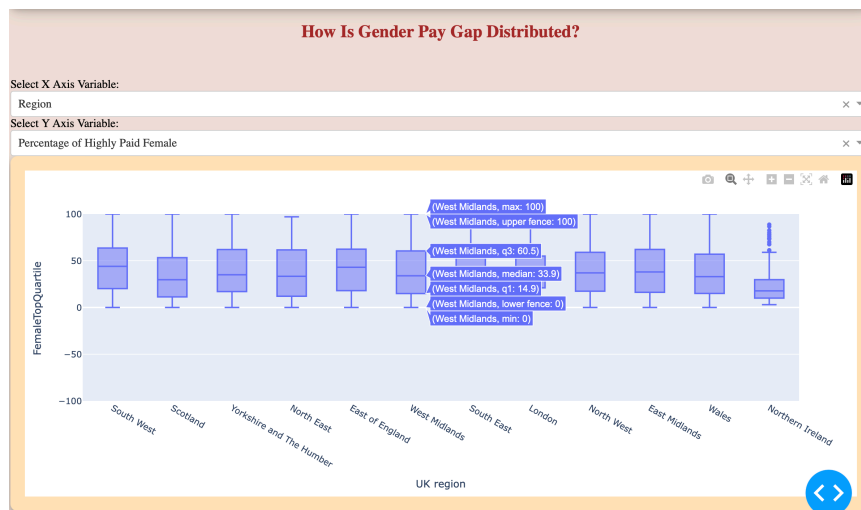- **How is my company's gender pay gap compared with others in the same dimension (industry, region, and employer size)? (For recruiters)**

Explain the design:

A Box and Whisker Plot (also know as box plot) summaries a set of data, which shows how data is distributed and any outliers. It is also useful when comparing different sets of data as multiple box plots can be drawn on the same axises. (ncl.ac.uk, undated) Since we aim to compare the distribution of different data in different dimensions, where each dimension includes various distinct categories, a box plot is quite appropriate due to its nature.

On our dashboard, users are able to view the results of gender pay difference and categories with median, quantiles, outliers when hover on the plot. We added functionality on top by providing users the selection of variable in both x-axis and y-axis based on users' interests and preferences. From x-axis, users can choose region/industry/company size in case of which from which dimension they want to compare. While y-axis offers choices among percentage difference if mean hourly pay, percentage difference of mean bonus pay, and top quantiles of female pay in case of which aspect they wish to make comparison. In total, there are 9 different plots.

Evaluate the design:

The box plot is useful with a quick and easy understanding of the overall structure of the data for users, especially female job applicants and recruiters. Users can view how large the size of gender pay gap is in different group, so as to decide the situation of a specific company in this regard. Besides, the box plot is efficient in dealing with large amount of dataset, as it allows for the compression of large amounts of data into a single, concise visual representation, making it easier to see the overall structure and distribution of the data. This means that users can still easily check the pay situation even with more companies in this system.

However, the box plot has limitation of comparing with more categories, e.g. both region and industry, because we only provide one y variable. So it may not provide a comprehensive view of all the factors if the user wishes to select all. Further, the choice of box plot on our dashboard will be complex for those without statistics background, especially if they are presented without clear explanations and labels. Hence, we can only answer question 1 by compare with target attributes respectively. We may incorporate a regression analysis or scatter plots in the future. These visualisations can help you identify specific patterns and trends in the data and provide a more complete understanding of the pay gap. In addition to this, we can provide a section with a short paragraph summarising the key features like median, quantiles, and outliers of box plots in words and figures.

3.3 Scatter Map
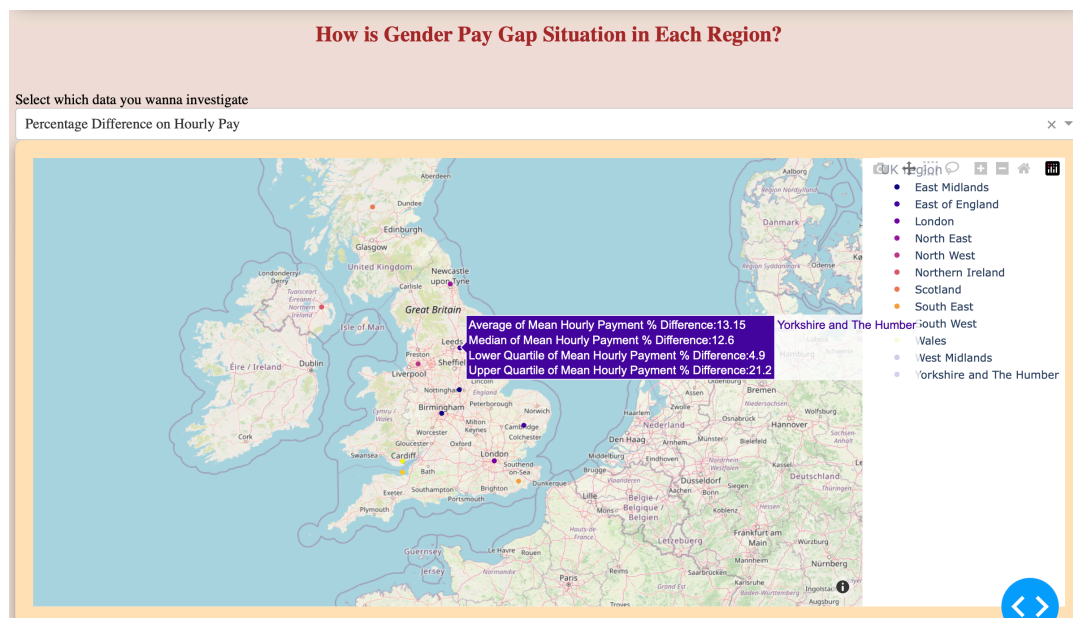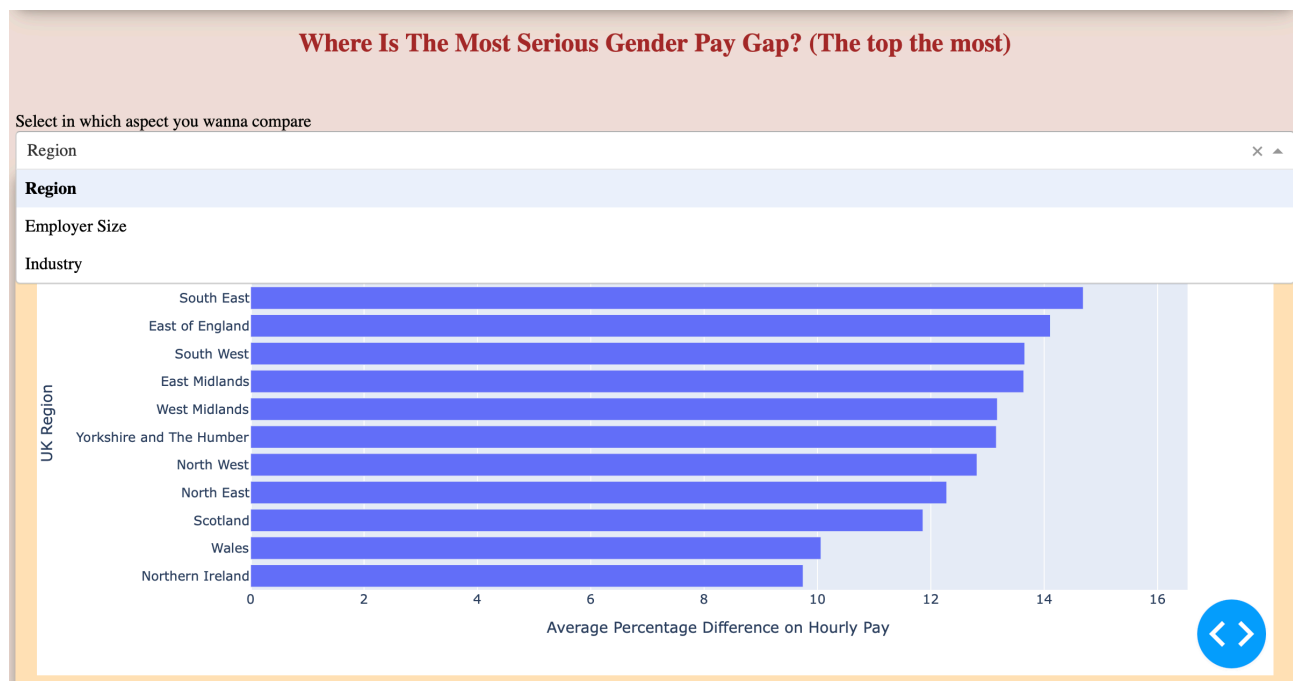
Explain the design:

In order to better present data in regional situations, we believe it would be best to show data on a map. Therefore, we have processed the original data by calculating the average, median, lower quartile, and upper quartile of Percentage Difference on Hourly Pay, Percentage Difference on Bonus Pay, and Percentage of Highly Paid Women across 12 regions in the UK and created three maps for users to choose based on their demands. When hover on the dot of a specific region, data from that region can be shown.

Evaluate the design:

No other types of plots can be more vivid than a real map in case of showing regional data. Meanwhile, as the data were summarised, users can easily get access to gender pay gap information of a specific region when they hover on that region's dot. However, the dot is not obvious. The ideal type of map is Choropleth map, which has distinct colours for the whole area of regions. Since we did not find an appropriate geojson to draw such graph, the dotted map is used as a substitute.

In addition, the map does not allow users to compare situations between different regions. It only shows information of a single region when hover on the graph and only one can be hovered at a time. It can be improved by adding a Box and Whisker Plot under the map, which saves corresponding regional data when the dot is on the map and users can choose which regions to keep by clicking on the legend of the map. Nevertheless, the coding capability of group members does not allow this to be achieved. Furthermore, the map can be zoomed in and out with Trackpad on a Mac laptop, while the graph takes a large proportion of the window, users may zoom the graph when they intend to scroll up or down. It would be better if the area shown on the map is fixed. Also, the map only contains 12 dots for 12 regions in the UK. In fact, we also have a column that specifies the address of each row in UK cities. However, it would be more complex to obtaining geographical coordinates of cities without a proper geojson. Though it would be more precise, plotting cities may make the graph in a mess with more dots and a separate dropbox should be required to select specific city to look at.

3.4 Bar Chart

Three bar charts are used to answer the question:

• **Which industry / employer size / region has the biggest gender pay gap?**

Explain the design:

When comparing metric values across various subgroups, a bar chart is an appropriate tool. This type of chart allows users to easily determine the most frequent or highest-ranking groups and compare them to others. (Yi, undated) Since we are about to rank and see which category in terms of 'Region', 'Employer Size', or 'Industry' has the most serious gender pay gap, bar chat is appropriate to show values of each category.

Evaluate the design:

The order of bars are ranked descending as the graph shows Average Percentage Difference on Hourly Pay of each group, where the larger number means larger gender payment gap. It provides a clear representation on which subgroup has the biggest gender pay gap in terms of hourly payment. However, since the data used to plot bar charts were created separately and each graph was plot separately with different datasets, if we want to make the comparison as what we did for the Box Plot, we will need 9 datasets and write similar codes for 9 times to draw the graphs. In case of future improvement, if not specified to use Plotly in this coursework, visualisation tools such as Tableau can be used to make it more convenient to draw different graphs.

Meanwhile, it also has limitations on comparing with more categories, e.g. both region and industry. Maybe the combine effect of multiple dimensions will be different with only looking at a single dimension. Hence, machine learning tools can be adopted to figure out how changes in multiple dimensions can impact on gender pay gap.

References:

Frost, J. (2021) *Using Histograms to Understand Your Data,* Statistics By Jim. Available at: https://statisticsbyjim.com/basics/histograms/ Accessed: February 2023

Mollerop, F. (June, 2019) *How The Gender Pay Gap Could Destroy Business Competitiveness.* Forbes. Available at: https://www.forbes.com/sites/forbestechcouncil/2019/07/26/how-the-gender-pay-gapcould-destroy-business-competitiveness/?sh=5fc6d60968dc Accessed: February 2023

ncl.ac.uk (undated) *Box and Whisker Plots*, ncl.ac.uk. Available at: https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/data-presentation/box-and-whisker-plots.html Accessed: February 2023

Yi, M. (undated) *A Complete Guide to Bar Charts,* Chartio. Available at: https://chartio.com/learn/charts/bar-chart-complete-guide/ Accessed: February 2023