INTRO TO PROGRAMMING II MIDTERM ASSIGNMENT

**Question 1**

**Outline your project and which extensions you are building on to the template. (400 words)**

I have chosen the data visualization app since I am planning to follow the Data Science track and I think it is a useful exercise in light of my future studies. There are other two reasons for my choice the first is due to my passion for visualization and infographics and the second is related to my job which is with an IT company in the data and analytics intelligence business.

I have chosen to create 3 additional extensions similar to the ones illustrated in week 9 and week 13.

I have decided to focus on the EU and OECD data available at: <http://www.oecd.org> and <https://ec.europa.eu/eurostat>

I wanted to create 2 extensions based on European Union Data statistics where I could compare some key metrics between European Countries and OECD.

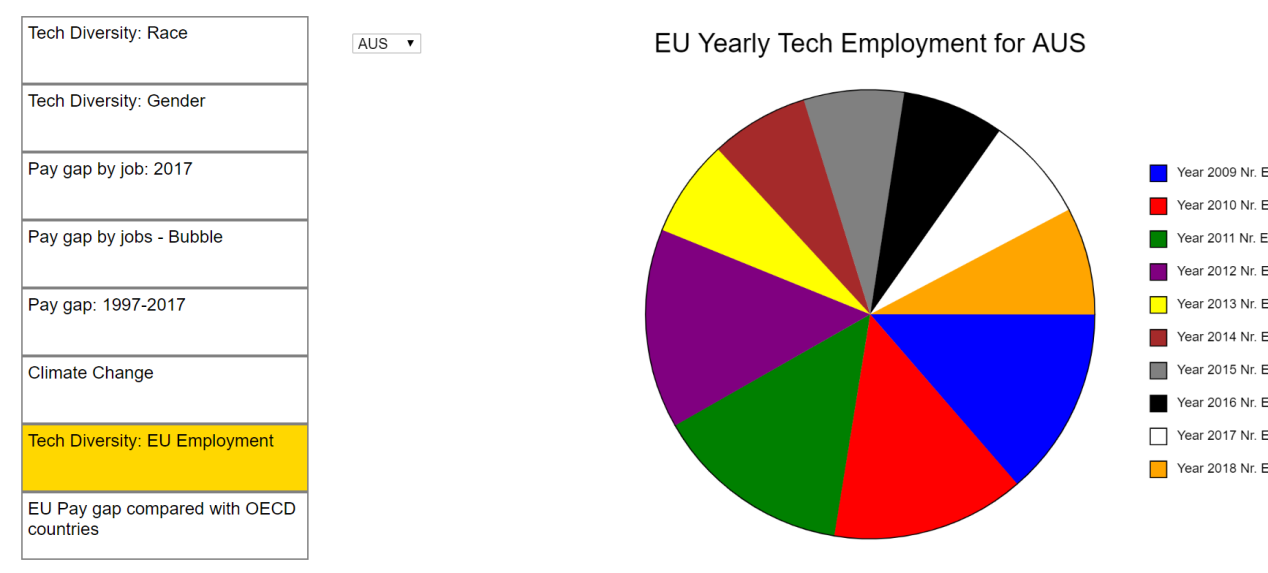
1. **The first extension Tech Diversity: EU Employment** is using a pie chart to illustrate the number of employees for each EU and OECD country in the last 10 years from 2009 to 2018.

You can select the country in the drop down list and get the number of employees for each year in the legend. I have reached a good progress in the implementation and I am planning to refine the visual aspects and data correlation in the following week to reach a satisfactory bug-free result.

1. **The second extension is EU Pay Gap Compared with OECD Countries.** The second extension that I have worked onto is a trend line again based on OECD Data at the same website as above and showing the yearly trend of wages in EU in comparison with OECD countries. I am using the historical data from 1990 to 2018. I am planning to refine the visualization and depnding on time available add further interactivity where possible.
2. For my **third extension** I want to create a **Pay Gap by Jobs Type Analysis ,** I am planning to use a bubble visualization and I am considering adding interactive features such as a slider moving from one year to the next and showing interactively the trend for each year. I am planning to use traversing of the database and row and column extractions either via nested loops or table operations that will allow for a maximum flexibility in accessing the data and visualizing it.

Timeline and outline of project is shown in GANTT Schedule.

The outline of the app menu as of this point can be seen below.



**Question 2**

**Discuss the progress you have made on your project so far. (400 words)**

I have researched the statistical websites: : <http://www.oecd.org> and <https://ec.europa.eu/eurostat> in order to understand what data is available and which data best fit my objectives. I have selected the csv data available on the websites and cleansed it using transformation processes and pivot tables to refine, filter and arrange the information according to objectives.

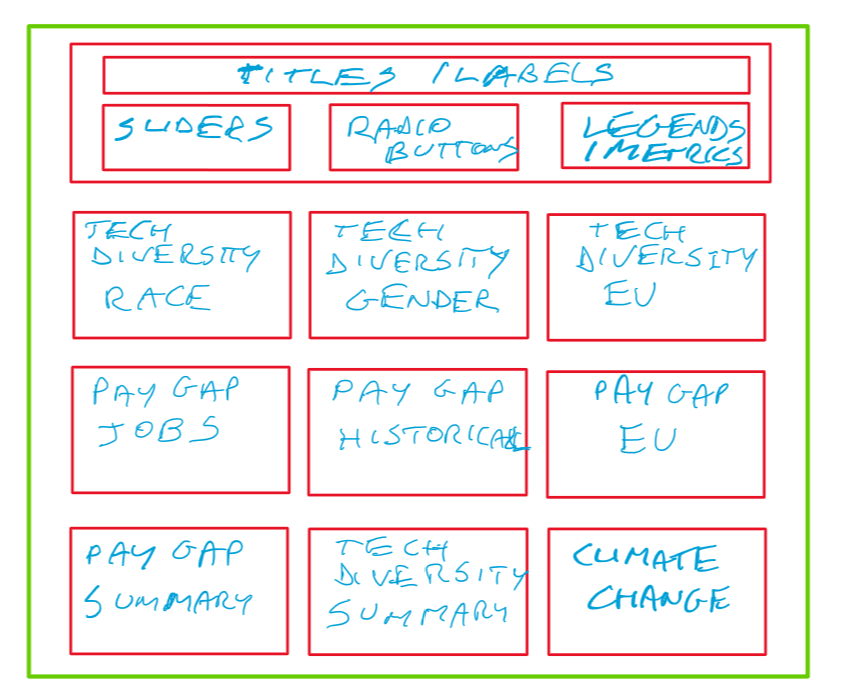
I am progressing rapidly toward completion of the first 2 extensions and I have completed the initial steps for the third extension which will be the focus of my job starting next week after midterm submission and all this in order to maintain a steady progress toward objectives.

I have written the core code for the first 2 extensions, but I have to complete testing/debugging and apply finishing touches. I am still facing an issue with a data column which is not being imported correctly and for which I am investigating solutions online on both Github and StackExchange.

The 3rd extension core code is in initial phase I have to investigate further a way to represent the selected data on the grid and I am looking at ways to add interactive elements such as hover-on legends, advanced metrics, interactive components, radio buttons and other visualization effects that update themselves dynamically. I plan to accelerate development of core components after mid-term submission is completed.

I feel I have a clear view of where I am heading, and I am committed to keep a steady progress toward milestones objectives. The outline/initial plan has been delineated for the project and I have a clear view of the steps and timeframes necessary to maintain consistency and meeting deadlines.

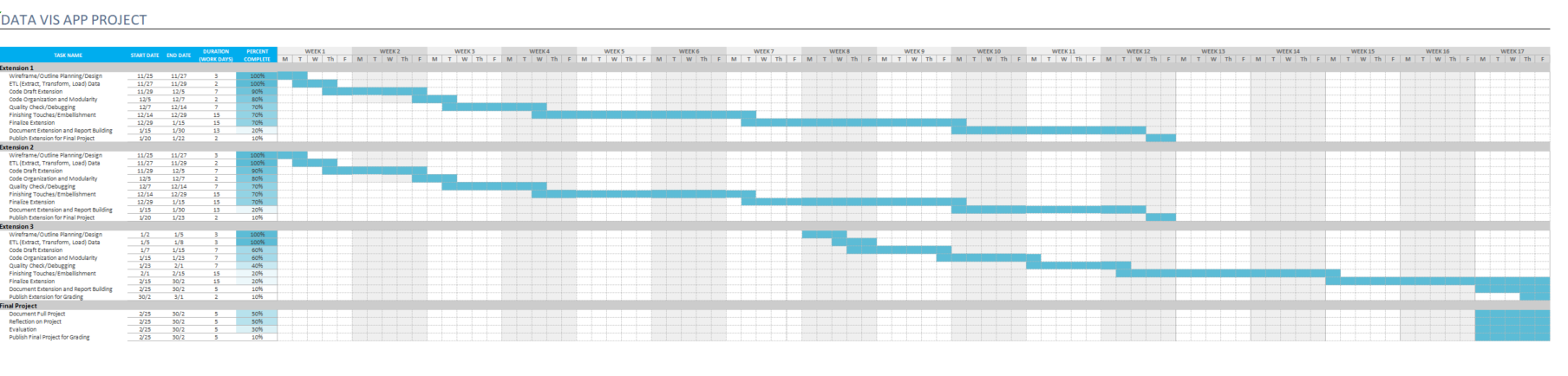
I am planning as an extra-mile aside from 3 required extensions to create a more evolved version of the app which will include a dashboard as home page for the app where the single extensions show as customizable widgets and queries can be compared and analysed from a single pane of view. The idea is to refine the app to higher standard and improve on project requirements although due to time constraints from my job I am leaving this out of the project schedule for the moment since I fear I could not be able to deliver it for the final project. This is something I am working onto but with a lower priority and based on time availability.

The wireframe of the dashboard I would like to create is the following: 

**Question 3**

**Discuss how you will organise you time for the rest of the project. (200 words)**

As outlined in the GANTT chart attached as Excel file in the “REPORT” folder, I am planning to wrap Extension 1 and 2 by January to concentrate my efforts on extrension 3 and allow ample time for revision, finishing touches, documentation and report building and avoid last minute glitches. I am planning to complete extension 3 by the end of the course as per deadline. Since it is more complex and demanding the the previous 2 extensions I want to devote more time to quality aspect to be safe that is meeting expectations. I have divided the time alloted to each extensions according to experience in developing the first extension and according to my strengths and weaknesses allowing more time to those parts I know by experience I will struggle with. Although I had also to consider integration of tasks, in the sense that I had to make sure to meet specific milestones which are necessary in order to move on to the next step, for example if the wireframe is not promptly ready and I do not have the big picture, ETL or Code Drafting cannot take place. I am making sure to adhere to the timeline.



**Question 4**

**List any external sources that you have actively utilised in your project.**

I have taken inspiration from a previous Coursera Specialization on Data Visualization I have done with Johns Hopkins to come with a design and for inspiration with data sources.

I use also extensively both Github and StackExchange together with Kaggle for inspiration and tips.