1. Track Selected: Cost of Living data analysis
2. Main Dataset: Cost of living increase in London
3. Track 1: *“Analyze the development of cost-of-living over time. Are there any detectable trends?”*

PART 1 – ANALYTICS

**a.**

1. Q1: *Analyze the development of cost-of-living over time. Are there any detectable trends?*
2. Q2: Changes in salaries compared to changes in cost of living, weekly, from the beginning of the 1st Lockdown to the end of the 2nd Lockdown.
3. Worry about the increase of living costs based on the Age, Social Grade, Ethnicity, Employment, Work Organisation, Gross Household Income and House Tenure.

The last section of the provided dataset will be used to answer to this question (more specifically PAGE 22-24).

**b.**   
**Q1**: In order to answer Q1, and analyse the development of cost-of-living over time, I will use the dataset present on the National Office of Statistics, released 15 September 2021, which contains all the data that I need to analyse the changes from Mar 2020 to July 2021. The data presented in the dataset is a time series, containing the index as well as the month, enabling easy plotting to visualize change.

The dataset is available at: <https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceinflation/current>

**Q2**: As per the information released by the Parliament, the date ranges from March 2020 (beginning of the first lockdown) to July 2021 (end of 2nd lockdown). The source is: <https://commonslibrary.parliament.uk/research-briefings/cbp-9068/>

The extra datasets used to answer this question are taken from the Office of National Statistics, available at the following link: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/averageweeklyearningsearn01/current:

* Dataset released on 14 September 2021 07:00
* Dataset released on 12 May 2021 10:41
* Dataset released on 10 November 2020 07:00

**Q3:** For Q3, I will be using the provided Dataset, which consists of a survey created in January 2022, from 1188 people, presenting a percentage split between different regions, age groups, household incomes etc. All these percentages can be useful in visualizing the split between everyone based on the aforementioned characteristics.

**c.**

The 5 datasets used, 1 for Q1, 3 for Q2 and 1 for Q3, are all focused on the financial aspects during the covid period. The first 4 assess the changes in salary as well as changes in customer prices index throughout the covid periods, meanwhile, the last dataset asserts the worry of the users that went through this covid period and financial changes and their sentiment.

PART 2 – DESIGN AND DISCUSSION

a.

Q1:

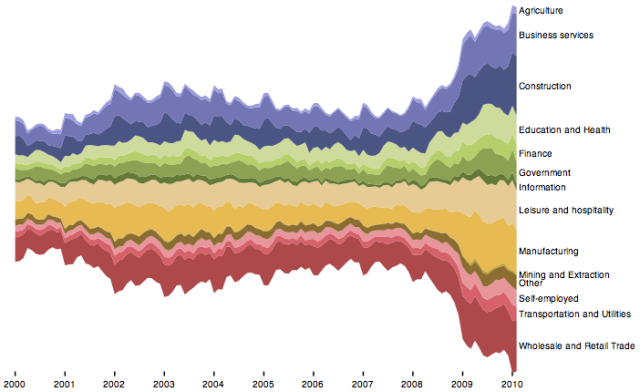
For overall view:

Image reference: <http://complexdatavisualized.com/time-series-visualizations-an-overview/>

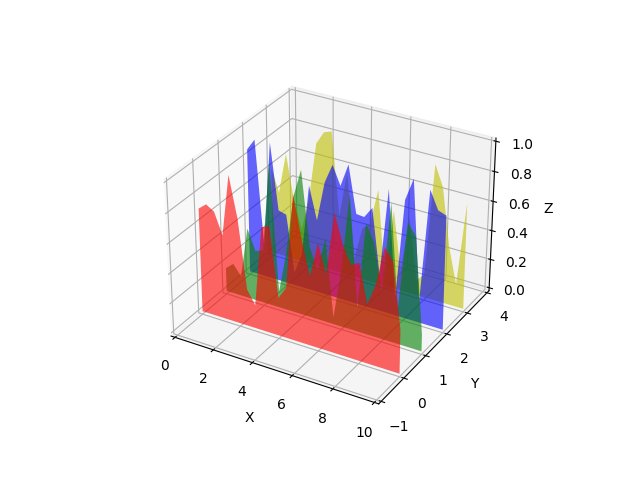
Y-Axis: TYPE

X-axis: time

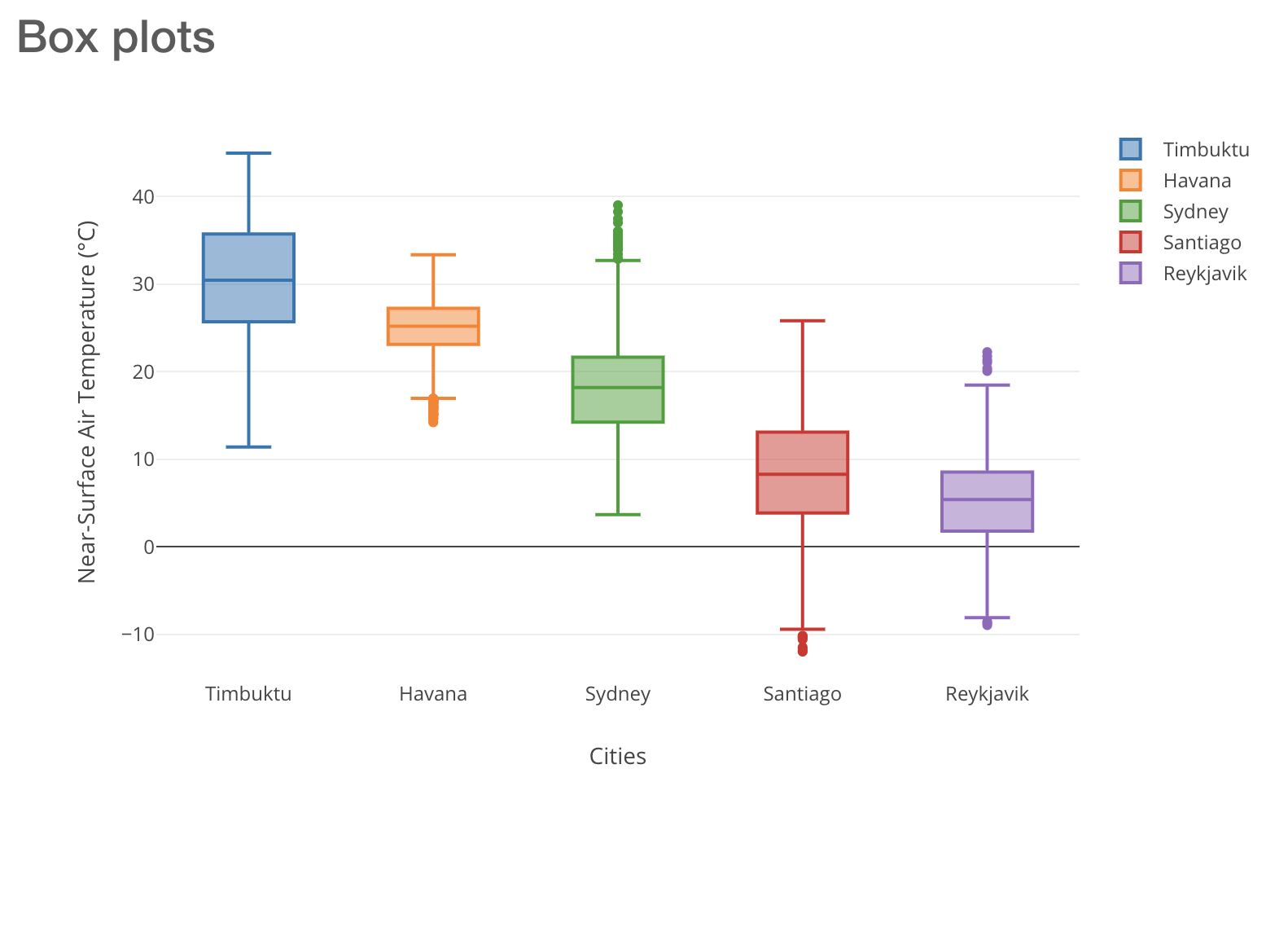
Different colours for different sectors

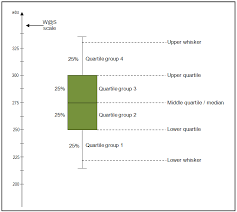


Then by clicking on one of the sectors, a 3D time-series appears:



Q2:





A box plot will be used to plot the time series data where the minimum value will represent the lowest salary industry and the highest value will present the highest value, and also the whole industry median value.

X-axis: time

Y-axis: value

Colour idk: to be seen later

Q3:

Multiple Sankey visualisations for all the different characteristics of the survey, with the right division, to form a dashboard.

