Data Visualisation 1 Report

FIT 3179 Data Visualisation
29802245 – Jeyvan Viriya
https://github.com/jeyvanVD/FIT3179
https://jeyvanvd.github.io/FIT3179/

1. Domain

The domain I have chosen for Data Visualisation 2 shows the comparison of living prices with comparison to New York, USA in the year 2020, where the chosen data and visualisation would be showing the users the state of prices in their country, with respect to some of the most basic parts of people's spending, including cost of living, cost of living plus rent, rent restaurant price, groceries, and the local purchasing power of each of the 132 countries given in the data.

The dataset was chosen from the year 2020 because from the year 2019 to 2020, there has been some fluctuations within the global economy caused by the COVID-19 pandemic, resulting in changes of prices like no one has ever seen before. Thus, making the visualisation fresh and factual, showing the world as it is right now.

2. Why?

The main goal of this visualisation is to give a clear overview on how cheap or expensive a user's basic spending is, in comparison to people in other countries all around the globe. Hopefully by doing this, the users could be more considerate to other countries whose state of economy is worse rather than the user's country.

The visualisation also aims to give the user additional information with regards to which countries have the highest and lowest amount of expenditures when considering living abroad.

3. Who?

The target of this visualisation would be the general public who wish to understand the financial state of their country and other countries around the world, and also answer the question "Is there a country more expensive than mine?". This visualisation can also be utilised by both governmental and non-governmental organisations as a source for research and basis for comparative study in the socio-economic structure of the world.

4. What?

The data is based on Numbeo, which is a crowd-sourced global database recording the quality of life for people globally, comprising over 6 million prices in 9 thousand cities all around the world (Numbeo, 2020). All the estimations that are made within the data are based on the comparison of prices in each country to that of New York City, USA.

5. How?

The visualisation starts by giving the user a global view of the world map in "equirectangular" projection This projection is chosen because it shows the entire map cleanly and clearly with a balance between angle and area. This map would be a choropleth maps because the purpose of the map is to show the categorical value of a fixed range between the different cost of living index among countries.

The cost of living index values are divided among 5 categories, each showing the countries whose cost of living index is between 0-25, 25-50, 50-75, 7-100, and those over 100 with respect to the values in New York City, USA.

The visualisation would then touch on each of the quantitative values restaurant price, cost of living index, local purchasing power, rent, and groceries index. This part of the visualisation is designed with respect to the local purchasing power, but with different quantities including the restaurant price, groceries price, and rent price. Due to the values all being quantitative values, the idioms used are bubble plot and contour plot. Bubble plot shows the relationship between 3 quantitative values and 1 qualitative value (continent) which are shown using the position, size, and colour channels, respectively. The contour plot is a unique plot that creates a heatmap out of the distribution created by each quantity, and this is used to show the relationship between 2 quantitative attributes and a single qualitative attribute continent. The contour plot is expected to help the users see the trend on which each relationship is expected to build on.

At the bottom of the visualisation, a parallel coordinate plot, which takes in all the quantitative attributes and shows the relationship between them for each of the category (which in this case is the qualitative attribute continent). Again, this plot is used to show the users a clear overview of the expenditure levels of their countries, among others.

The unique attribute from the visualisation includes the contour chart and the parallel coordinate plot, both being rarely used and showing a clear overview of the bigger picture on the expenditure levels of the different countries.

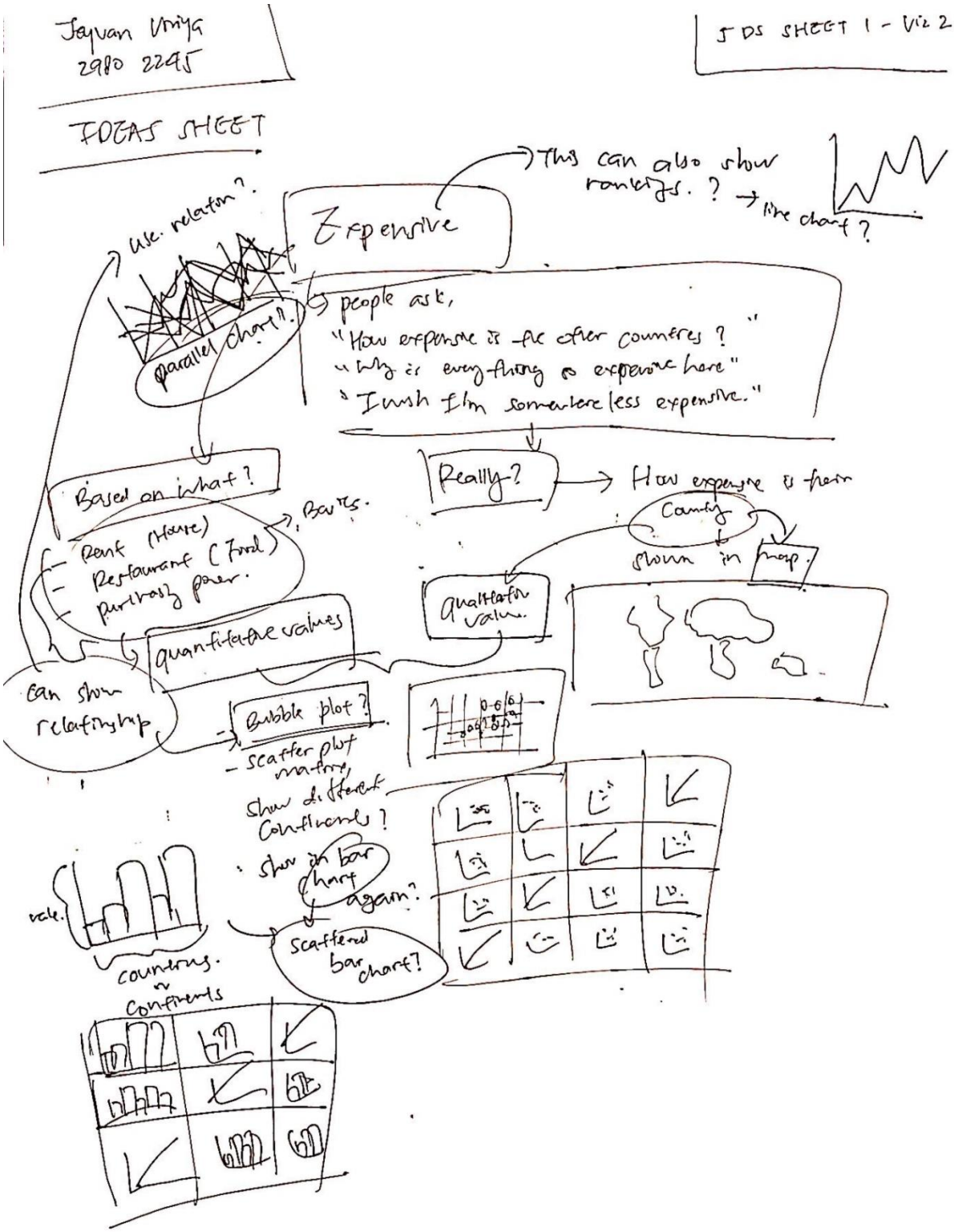
Bibliography

Cost of Living. (n.d.). Retrieved November 02, 2020, from https://www.numbeo.com/cost-of-living/

Cost of Living. (n.d.). Retrieved November 02, 2020, from https://www.numbeo.com/cost-of-living/rankings_by_country.jsp?title=2020

Yadav, V. (2020, March 21). Countries Dataset 2020. Retrieved November 02, 2020, from https://www.kaggle.com/dumbgeek/countries-dataset-2020?select=Cost+of+living+index+by+country+2020.csv

5 – Design Sheet



Isyvan Viriya
24/02/2024

SDS Sheet 2 - Viz 2

To show the map → Countries

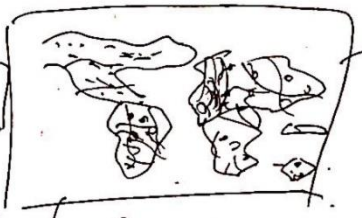
Cost of Living Index

Shows:
- Country
- Cost of living index

We have: qualitative. → Dot maps?

But we are representing whole countries

If we use dot maps, we can only show points.



too many dots → makes the user tired.

Dot maps.
→ But covers on the missing values.

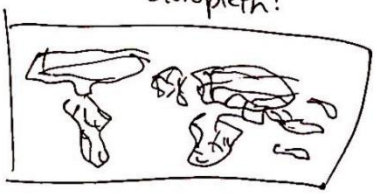
→ proportional symbol?
— same issues with dot maps.

→ Convert into qualitative?
— use limits. Let's say 0 - 120, then we divide.

0 - 20
20 - 40
40 - 60
60 - 80
80 - 100
100

Now we have 2 qualitative:
- Category
- Colouring.

→ Choropleth?



→ way cleaner
→ can show countries directly

↳ problem:
- missing countries?

Jeyan Viriya
2980 22415

ES DS SHEET 3- VIZ 2

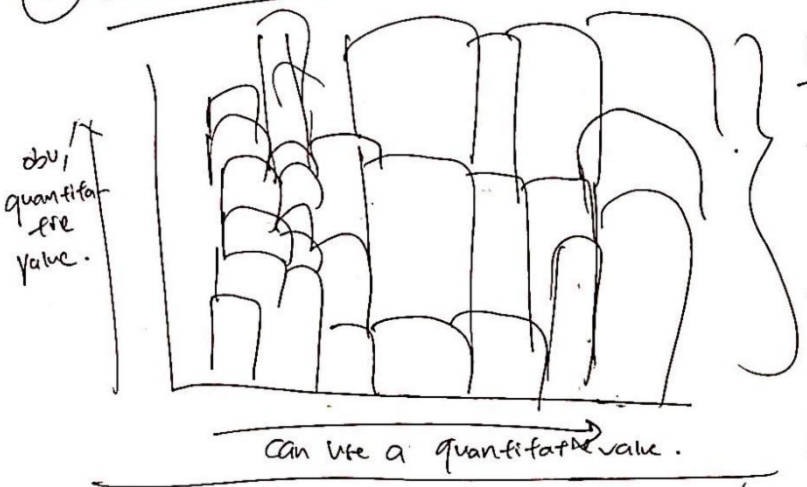
For the whole data

↓

all qualitative.

If show general values → all data (multiple quantitative).

(1) Stacked Bar chart.



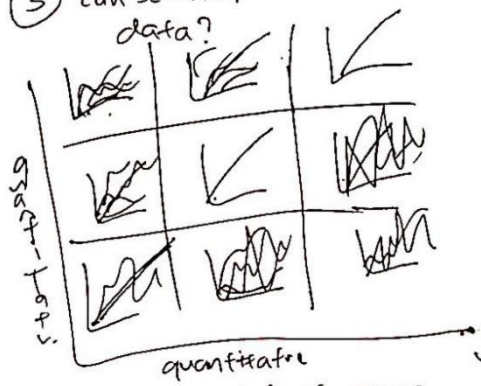
problem:
- messy when too many countries.
- does not look appealing to the eyes.

(2) Bar scatter plot

Instead,

Bar scatterplot?

(3) can scatter, but more data?



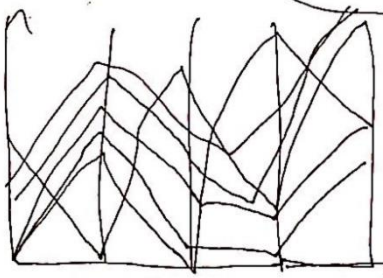
- also takes lots of space
- can show everything.
- can be hard to see values. Maybe just for trend?
- BETTER THAN BAR

- way more ink
- heat

- with many data
- can be messy.
- takes lots of space.

take category or continents → less messy

(4)



does not take space

can be messy

shows overall

parallel coordinate.
(shows all the quantities.

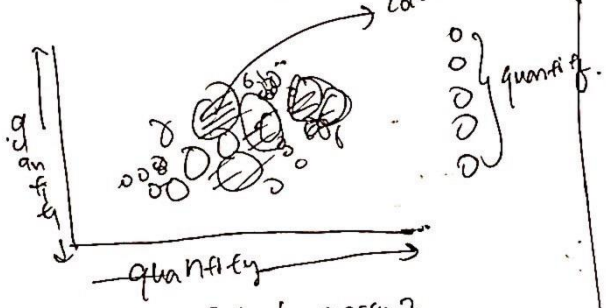
Jeyuan Viriya
2980 22645

IDS SHEET 4 - UE 2

For the data for each "category"

* Need to show all countries for user to see.
  must not be messy.
* show as much quantity as possible.

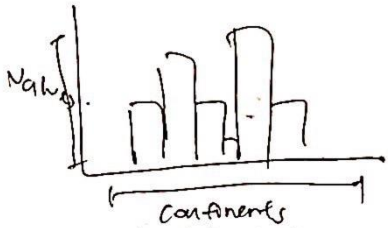
① Bubble plot
- shows 3 quantity 1 quality ✓



* but can this be messy?

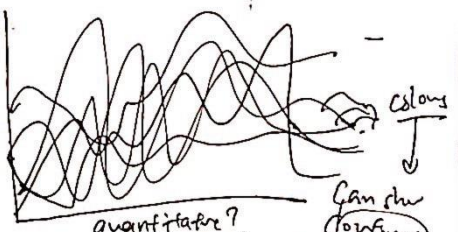
② Bar chart
- USED BEFORE
- hard to show countries. (How? Countries?)

Can only show Continents

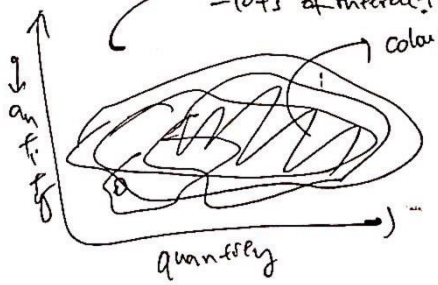


- very limited data.

③ line chart → also used before.



④ Contour plot?
- something new. What is this?
- can be compli cated.



Show country & continent side by side?

Jeyvan Vinaya
2930 7245

SDS- SHEETS VIZ 2

Final look?

- must be clean
- must show
  - map → choropleth
  - per county → bubble plot
  - per continent → contour
  - per "category" → parallel coordinate
  - all unique

Format