Ragger Jonkers - 10542604
Ellen van ‘t Klooster - 10207309
Belle Bruinsma - 10676759
Urscha Fajdiga - 11377437

Week 1

Analysis:

- What trends do you see in the data?
  - From 1950 to 2015 linear trend.
  - Same for 1950-2050
- Analyze how big the differences between various estimates are. Do you see a trend, i.e., do the differences become smaller or larger over time?
  - The differences in the estimates become larger over time. This is because the population is estimated and extrapolated.
- Think about these differences relative to the estimates at the respective time points and in absolute terms. When are the uncertainties the largest in absolute, when in relative terms?
  - The further in time, the more uncertain the estimates become in absolute numbers.
  - In relative terms, the difference becomes larger as well. (2% difference in 2030 between the first and second estimate compared to 4% in 2050)
- Do you think you can faithfully represent the uncertainty and the data in the same plot? Why, or why not?
  - Yes you can represent the data in the same plot. You can show the uncertainty with a dotted line.
- What effect do you think will the linear interpolation have on the uncertainty?
  - It won’t make the uncertainty bigger because the interpolation is the most logical choice for this dataset.
- Is linear interpolation a suitable method for this data?
  - Yes because there are no extreme outliers in the dataset.

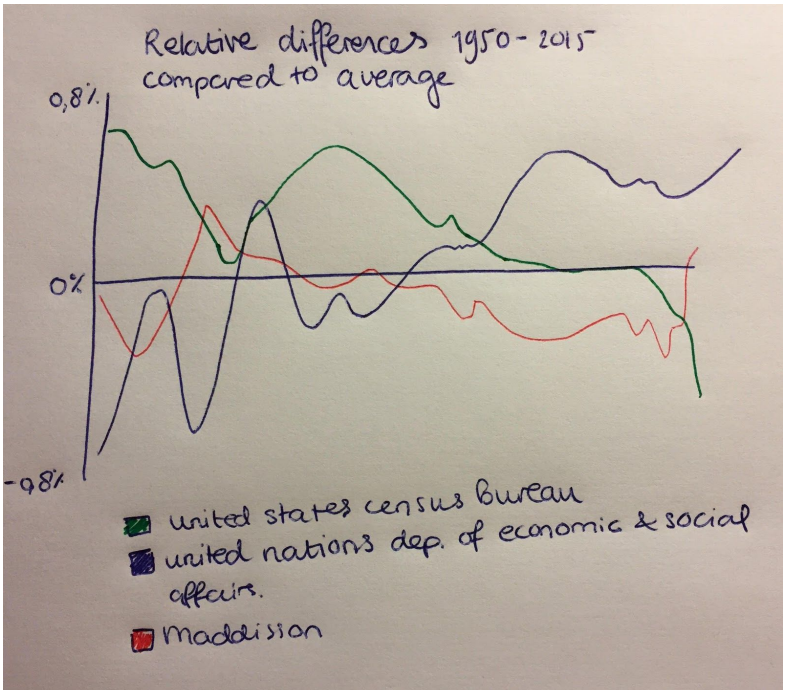
Sketching:

Design four alternative visual representations for representing the data and the uncertainty in the data. Consider different scales, both for time and for population numbers. You should design for an interactive system, i.e., you should not assume that you have to fit all content onto paper. Please take the theory from the video lecture on Graphic design into account. Here are some points you should consider:

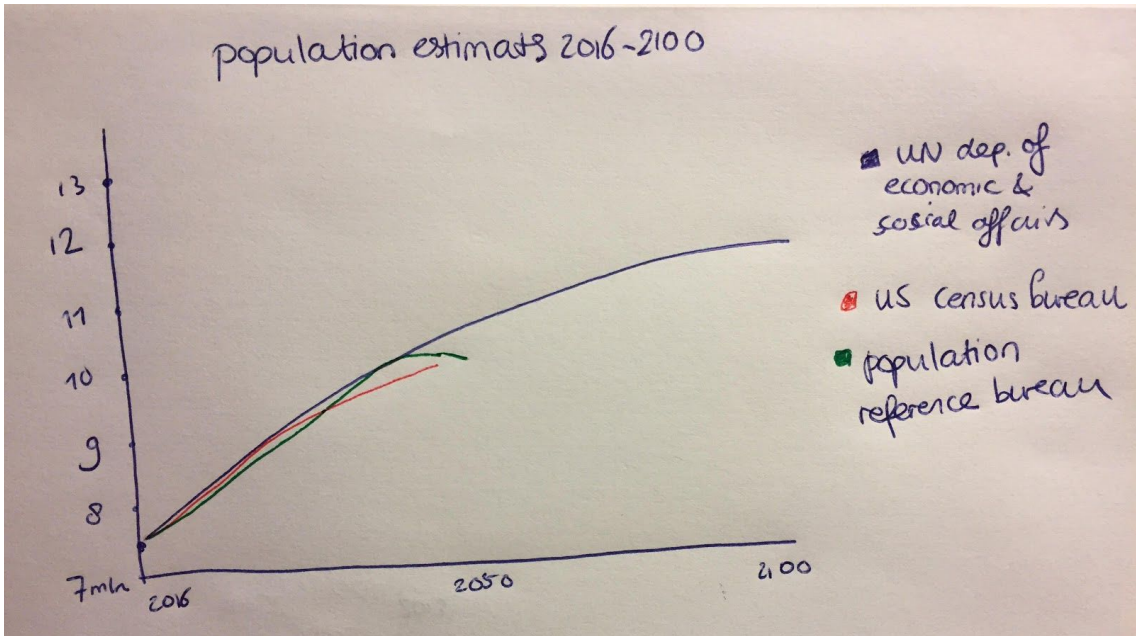
- To get a feeling for the final visualization, try to draw the data to scale.
- Instead of or in addition to showing 5 conflicting lines, develop a visualization that shows the data and the ambiguity. You can use a single visualization, or you can use multiple views.
- Your visualization should show the divergence between estimates in absolute terms (i.e., the difference in number of people) as well as in relative terms (i.e., % of divergence/uncertainty relative to a consensus value for a given year).
- Your visualization should make it easy to read a specific “consensus number” for every year.

<u>Visualizations:</u>

1.This graph shows the relative differences compared to average 1950-2015. The consensus number here is the average world population of the three different bureaus. The relative differences between the bureau and this number are shown in the graph. This difference is never above or below 1% as the graph shows.

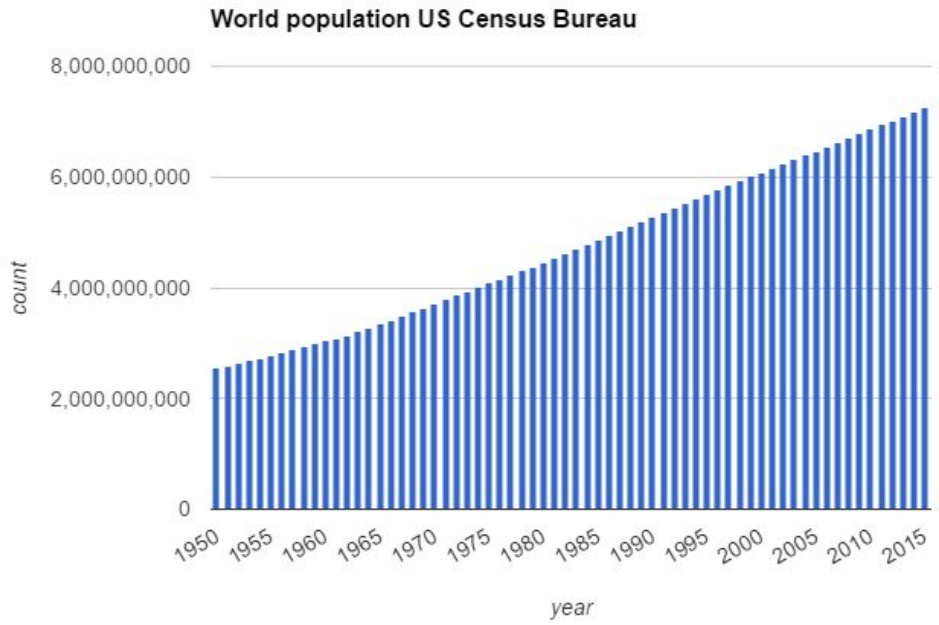


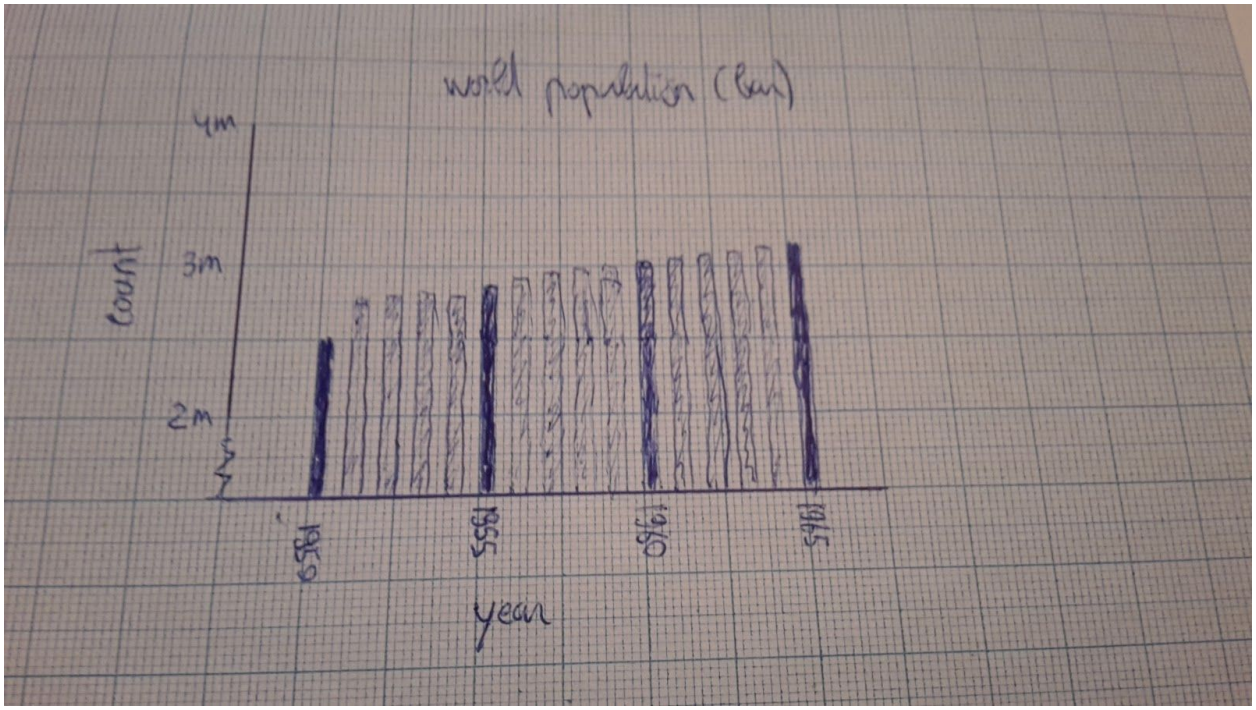
2.Population estimates 2016-2100. This graph shows the forecast of population as by three different bureaus.



3. For this visualization, the average of the two bureaus with the most data was used again (between 1950 and 2016). This was the United Nations Department of Economic and Social Affairs and the United States Census Bureau. The annual increase of the average of the two bureaus is shown in relative and absolute terms. As you can see, there is an absolute increase in the world population but a the relative numbers show a decrease.

4.It would probably be easier to read when split across multiple graphs covering these 65 years in 3 to 4 graphs. Another solution would be to color every 5th year differently, or at least easily recognizable. These solutions are shown in the (first of the possible 3 or 4) hand drawn bar graph below this graph.





Group reflection

We agree that visualization 3, which is the one that shows the changes in world population in numbers as well as relative percentages is the most useful.
This graph shows the most data, since it combines relative and absolute increases and decreases over the years. We think that looking for information about world population, this might be the most effective way to show information. The reason for this is that the data is likely wanted as a mean to observe either decrease or increase in world population rather than a specific value for a year.
Also it solves the scaling problem: it can be hard to read differences when the y-axis stands for world population count, instead of the percentages in this graph in which variation is easily perceived.