Students:
Maarten Hogeweij
Martijn Kruiten
Chris Olberts
Sanne Strikkers

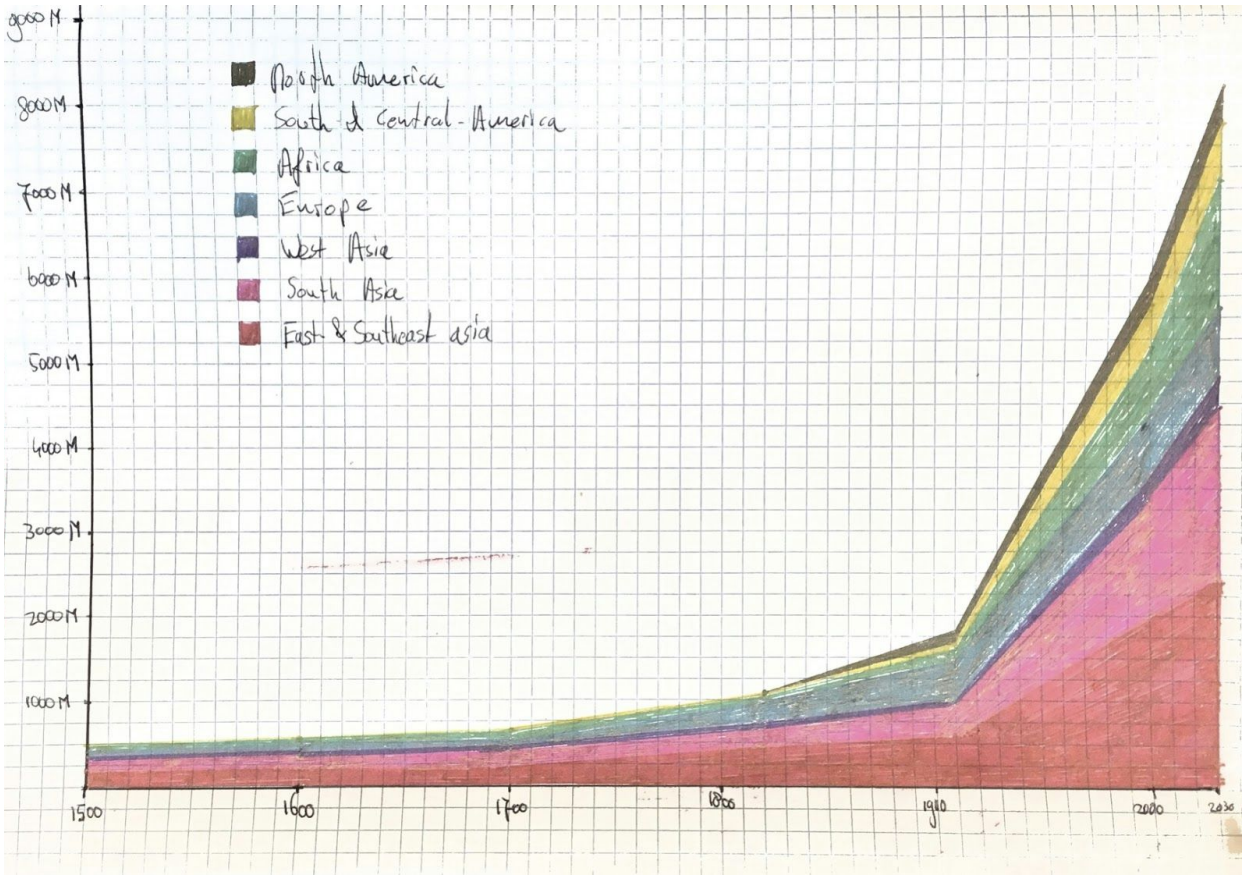
PART 1 - ANALYSIS

- What trends do you see in the data?
    - " Exponential" Growth: Slow at the start, fast grow at the end.
    - In the future the growth will be slower than the growth from 1950 to 2015.
- 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 relative differences become smaller over time. The estimates become more specific over the time.
    - Also the data between the groups are less different. So the techniques that the groups are using become better over the time and the measurements will be more specific. The amounts of the different groups are closer to each other.
- 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?
    - Before 1950 they are much larger in relative terms, after 1950 in absolute terms.
- Do you think you can faithfully represent the uncertainty and the data in the same plot? Why, or why not?
    - If you plot population against years, you can use error bars to represent uncertainty. Larger error bars indicate a greater uncertainty. A suitable way to give the error bars a representative length would be the standard deviation.
- What effect do you think will the linear interpolation have on the uncertainty?
    - Before 1950 the growth is exponential, so the error will increase with linear interpolation. After 1950 the growth is more linear and based on this data alone, a linear interpolation will give a proper estimate.
- Is linear interpolation a suitable method for this data?
    - After 1950 the data is more or less linear. Based on this data alone, a linear interpolation over a short period can be a reasonable estimate.

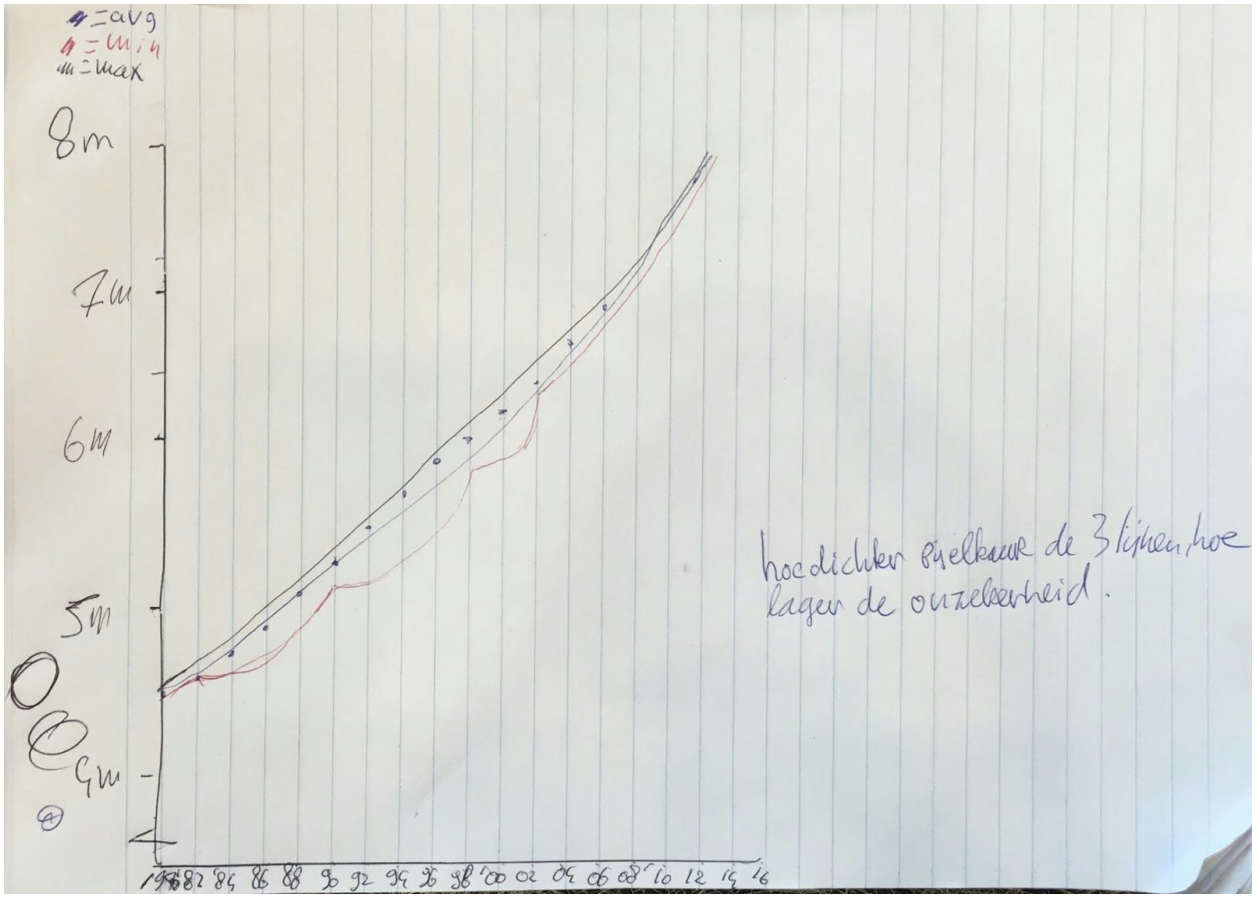
PART 2 - SKETCHING

Options

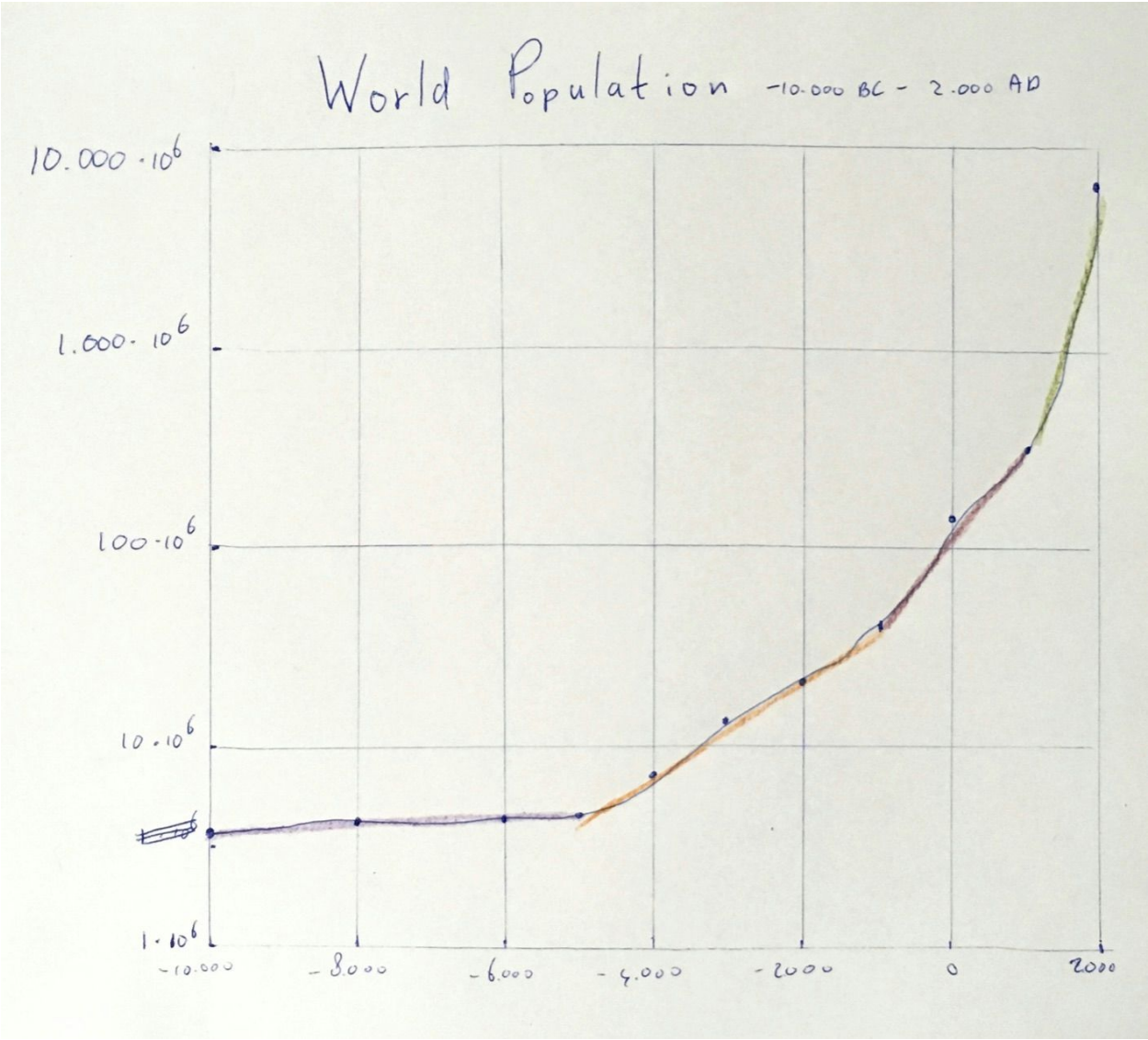
1. This graph shows the world population of different regions of the world from 1500 until 2030 (prognosis). The X-axis displays the year and the Y-axis displays the population size in millions (M). The colours represent different continents/regions as indicated in the legend. The stacked data is a representation of the total world population. Measurements and estimations were done for the years 1500, 1600, 1700, 1820, 1913, 2000 and 2030.



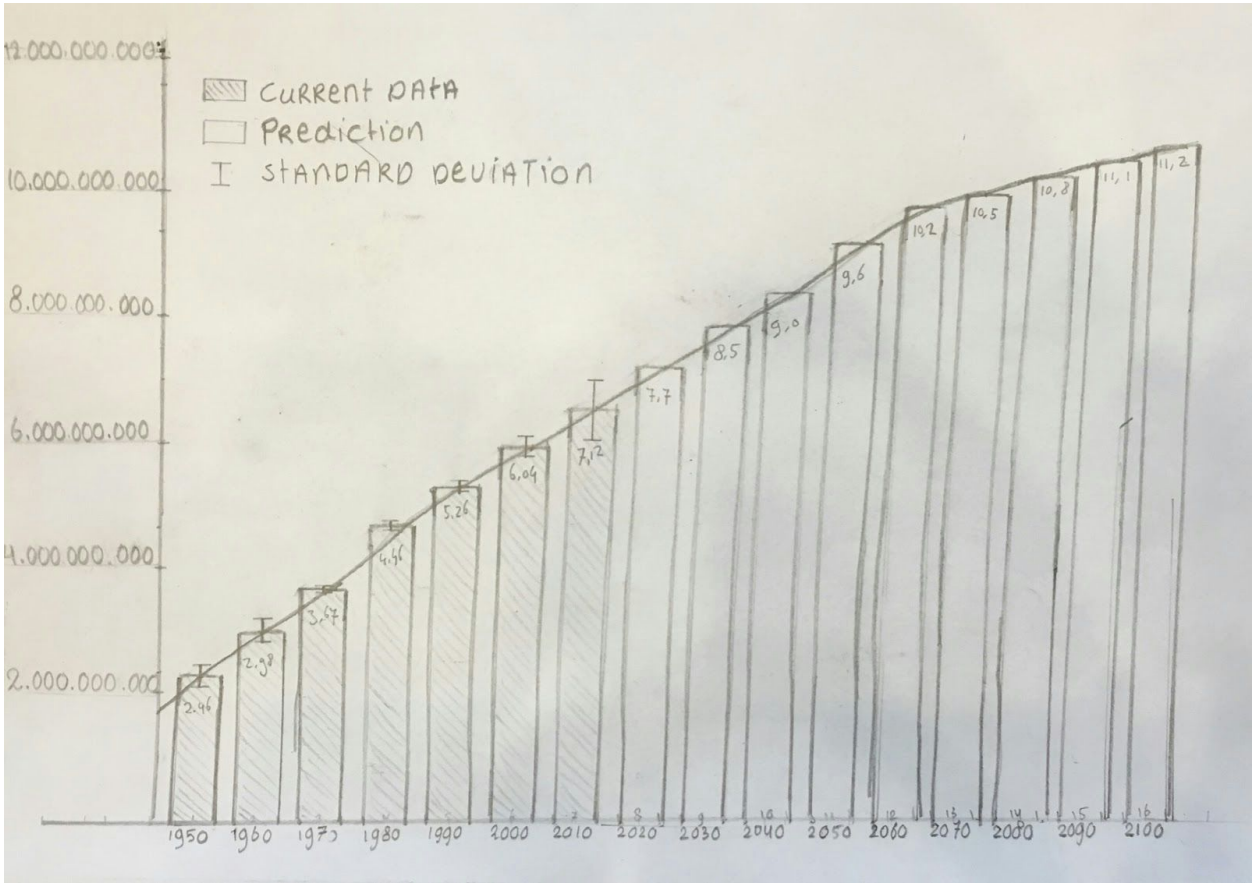
2. This graph shows the data according to five different bureaus. The blue line represents the average between the bureaus. The red line shows the lowest estimate of the five. The black line shows the highest estimate. The closer the three lines are, the less ambiguncertainty uity there is between the estimates.



3. This visualization show the world population from 10.000 BC to 2000 AD. The scale is logarithmic, because otherwise it would be flat for anything up to relatively recently and extremely steep after that.



4. This graph shows through every ten years the population in milliards. From 1950 to 2100.The data is an average of the population taken from the several groups. Also the graph shows the current population and a prediction of the population. Then you will see that the growth is becoming less. The line makes the growth of the population clear. The vertical lines in the current population shows the standard deviation. Also called the error bars.



PART 3 - GROUP REFLECTION

1.
    a. Advantages
        i. The stacked layout of this graph not only show the population of each different part of the world. but the top of the coloured area represents the world population from 1500 to modern times (and a little prognosis for the coming years until 2030).
        ii. One can clearly see the current population distribution over the world as it is now.
        iii. One can also sort the data in groups of own interest (i.e. Eurasia, America).
        iv. Historical events are visible: Colonisation of America, Industrial revolution in Europe, etc.
    b. Disadvantages
        i. There is no inaccuracy displayed over the total world population or the individual continents
        ii. The colours are not well chosen. The contrast between neighbouring groups should be higher to make a better distinction where the line ends.
        iii. It only displays data from 1500 to 2030

2.
    a. Advantages
        i. The graph shows the growth of the world population as an average between different census bureaus, as well as the ambiguity between different census bureaus, in the same visualisation. Because of the small scale the graph is quite precise.
    b. Disadvantages
        i. The graph does not show the data according to a single bureau, but instead shows the average. If you want to know the data from one bureau, the graph is useless.
        ii. The small scale makes it impossible to look back far in history. The graph also lacks projections about the future.
        iii. The graph has been sketched very roughly.

3.
    a. Advantages
        i. Differences are easier to distinguish for data before 1900 when compared to a linear scale.
        ii. Different growth rates in different times are easy to distinguish. They can probably be linked to major historic events.
        iii. It allows for taking large time spans without having a flat line for most of the years and an extremely steep line for the more recent years.
    b. Disadvantages
        i. The logarithmic scale makes it hard to visually compare different years. The Y-axis needs to be very detailed and one would have to look up different values and remember them in order to compare the two.
        ii. Error bars would have no meaning on a logarithmic scale, because they are way too hard to interpret.
4.
    a. Advantages
        i. The error bars make it easy to see uncertainties in data over the years.
        ii. The linear scale and the fact that the values are in the same order of magnitude make it easy to compare different years.
        iii. The reader will see an average of the population through every ten years. Enough to see the growth between every ten years.
    b. Disadvantages
        i. The visualisation could be better. It is not so clear for the reader what is happening in the graph.
        ii. Also the graph is not interesting to look at.

Discussion

We decided to pick graph 4 as our main pick because this graph best reflects the assignment goals.

The graph shows both the consensus data and the ambiguity. Absolute ambiguity is drawn in the shape of error bars displaying the standard deviation. From the size of the bars is immediately apparent that the population growth rate declines in the future. The bars show the trend at glance, but when one wants to see more detailed data, the numbers are written to give precise information on the population size. It also has a trend line which makes it easy to read the numbers in between. It shows the absolute ambiguity, but it’s also possible to estimate the relative ambiguity, by comparing the error bars to the population bars.