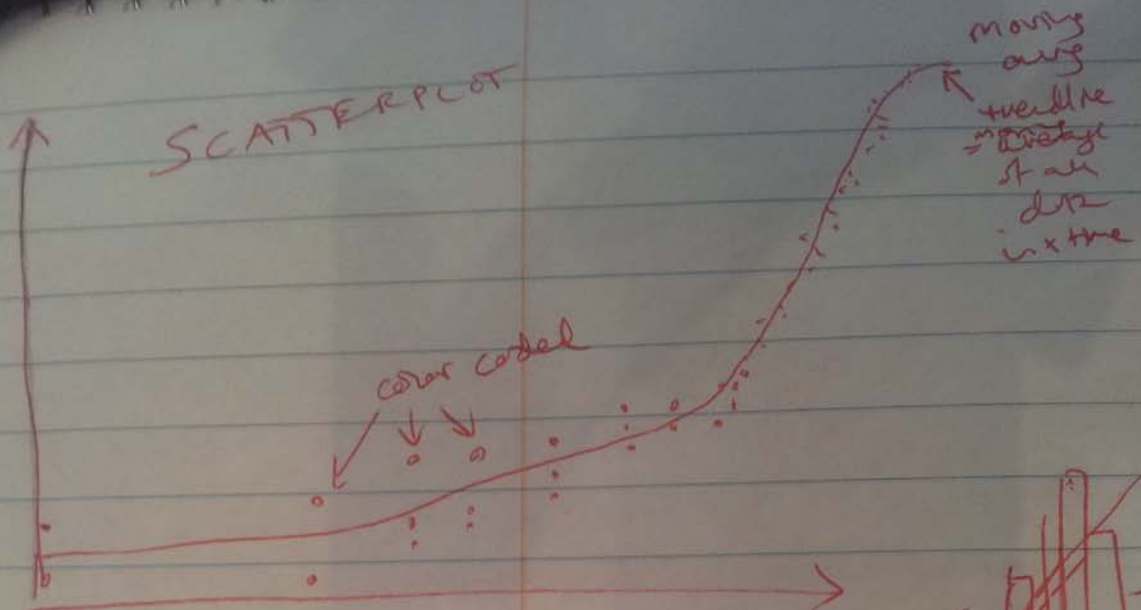


Part 1 – Analysis & Designs

- What trends do you see in the data?
 - **There is relatively very little change in the earliest time periods, for several millennia, after which the rate of change increases gradually.**
 - **The population growth is exponential.**
 - **There is more data available for current and recent time periods.**
 - **Currently, we are adding between 300M to 400M people to the planet every 5 years.**
- 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 estimates vary wildly in the earliest eras. For example, Thomlinson estimates that in 5000BC, the population is anywhere between 5M to 20M. In 0AD, the Population Bureau cites 300M while Tanton cites 150M.**
 - **Over time, these become smaller. In 2005, the maximum variation amongst the 4 estimating agencies which provided data for that year is only ~56M, which is less than 1% variation from any given estimate for that year.**
- 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?
 - **Using Excel, one can determine that the largest absolute uncertainty is in 2020, with a difference of 442M between the largest and smallest data (column 1 vs column 6). The largest relative variation, counting the mean of all available data (including ranges “x to y”), is in 5000BC, with a variation of almost 85% from the mean (min 5,000,000; mid 12,500,000; high 20,000,000)**
 -
- Do you think you can faithfully represent the uncertainty and the data in the same plot? Why, or why not?
 - **Yes, it is possible, for example by using a dual overview/detail arrangement of line charts, such as is in the UN Health graphs.**
 - **By doing so, we can separate the uncertainty from the data in the same visualization.**
 - **We are also providing more detail to the user at any given point.**
 - **Due to the complexity of the resulting dataset, interactivity may have to be used to make it easier for the user to use the graph.**
- What effect do you think will the linear interpolation have on the uncertainty?
- Is linear interpolation a suitable method for this data?
 - **Linear interpolation is not an accurate estimate of the population because the slope is exponential. Therefore, any linear estimate will either overestimate or underestimate the real curve. Hence, Linear interpolation is not suitable for this data.**

SCATTERPLOT



bar charts +
trendline
(but messy!)

color coded columns

