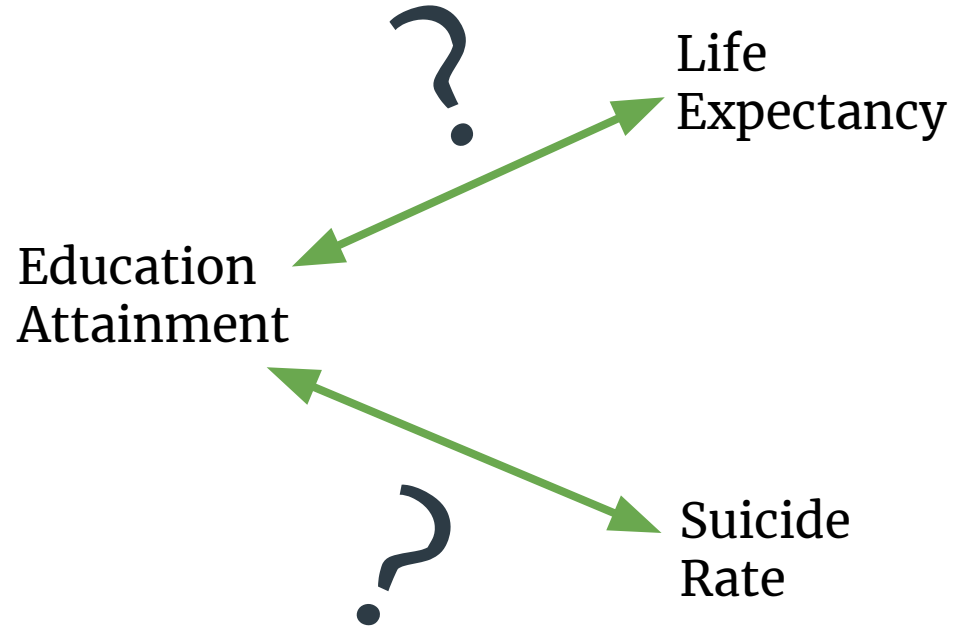# Project Edu +

Group #5

Amitabh Nag
Anqi Cao
Shiyao Chen
Susan Zheng

Agenda

- Project Motivation
- Project Goals
- Data Overview
- Data Sample & Cleanup
- Related Work
- Prototyping Process
- Draft Dashboard
- Usability Test
- Demo

Project Motivation

Life
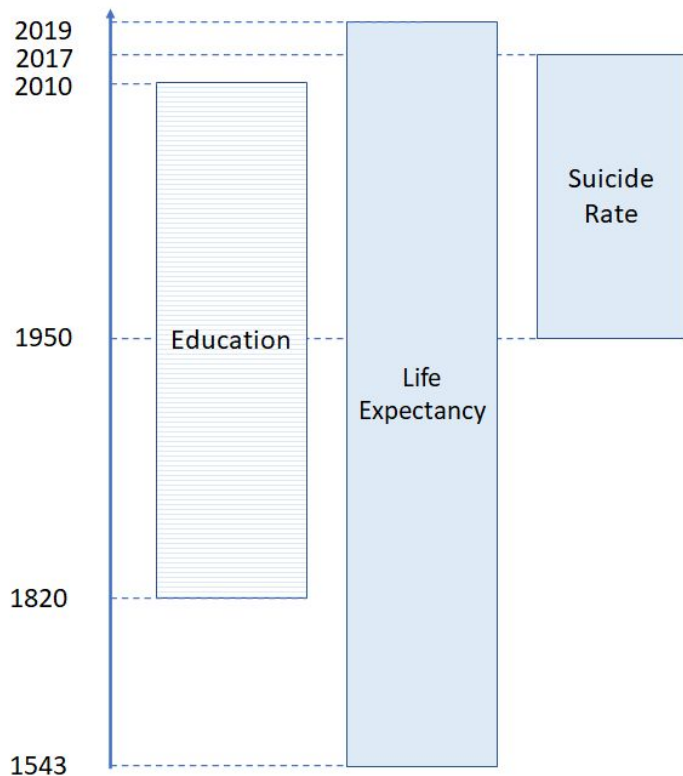Expectancy

Education
Attainment

Suicide
Rate

Project Goals

Questions/Hypothesis:

1. Is education attainment increasing over the years?
2. Asia-Pacific, Africa and Latin America have higher rates of changes in education attainment compared to Advanced economies (North America, West europe etc)
3. Is there a correlation between average life expectancy and education attainment?
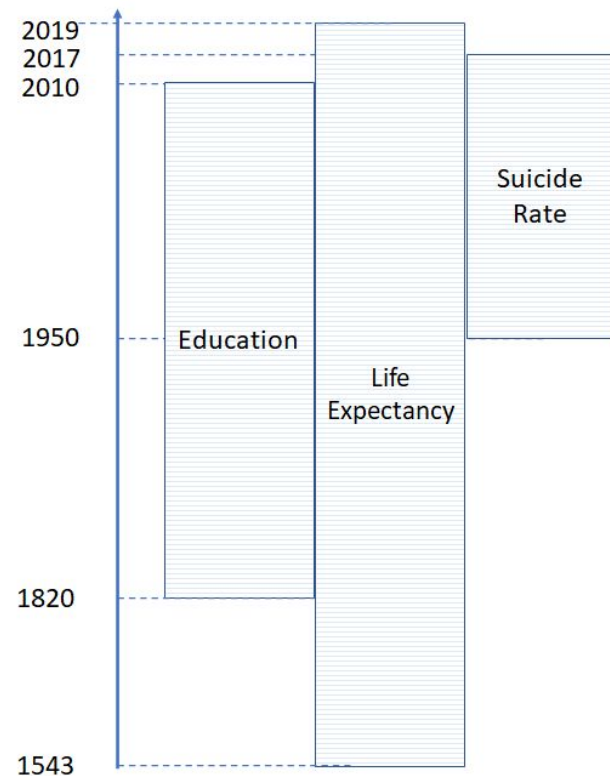4. Is there a correlation between suicide rate and education attainment?

Data Overview

- Three data sets:
  - **Primary: Long-term education attainment**
    - Education attainment from 1820 to 2010 by sex; 5-year intervals and country
    - 12987 rows x 23 columns
    - Variable types: 20 ratio; 2 nominal; 1 interval
    - http://www.barrolee.com/
  - **Secondary: Life expectancy**
    - Life expectancy from 1543 to 2019 by country
    - 19028 rows x 4 columns
    - Variable types: 2 nominal; 1 interval; 1 ratio
    - https://ourworldindata.org/life-expectancy
  - **Secondary: Suicide rate**
    - % share of deaths from suicide from 1950 to 2017 by country
    - 6469 rows x 4 columns
    - Variable types: 2 nominal; 1 interval; 1 ratio
    - https://ourworldindata.org/suicide

Data cleanup – Phase 1 (python)

2019
2017
2010

Education

Life Expectancy

Suicide Rate

1950

1820

1543

Python merge

2019
2017
2010

Education

Life Expectancy

Suicide Rate

1950

1820

1543

Cont.

Data cleanup – Phase 1 (python)

After some analysis, found that below countries in suicide/life-expectancy are similar, but different than barro lee

'Congo, D.R.','Cote dIvoire','Dominican Rep.','Hong Kong, China','Republic of Korea','Reunion','Russian Federation','USA'

Renaming these in barrolee data

```
In [68]:  #Country cleanup in barrolee
          barro_lee_data.loc[(barro_lee_data.country == 'Congo, D.R.'),'country']= 'Democratic Republic of Congo'
          barro_lee_data.loc[(barro_lee_data.country == 'Cote dIvoire'),'country']= 'Cote d\'Ivoire'
          barro_lee_data.loc[(barro_lee_data.country == 'Dominican Rep.'),'country']= 'Dominican Republic'
          barro_lee_data.loc[(barro_lee_data.country == 'Hong Kong, China'),'country']= 'Hong Kong'
          barro_lee_data.loc[(barro_lee_data.country == 'Republic of Korea'),'country']= 'South Korea'
          #assumption based on wikipedia article: https://en.wikipedia.org/wiki/South_Korea
          #"Reunion" and "Hong Kong" is not present at all in suicide_death_rate_countries
          barro_lee_data.loc[(barro_lee_data.country == 'Russian Federation'),'country']= 'Russia'
          barro_lee_data.loc[(barro_lee_data.country == 'USA'),'country']= 'United States'
```

Renaming a very long column name in suicide data to a simple name

```
In [69]:  suicide_death_rate_data.rename(columns = {'Deaths - Self-harm - Sex: Both - Age: Age-standardized (Rate) (deaths per 100,000 indi
```

First merge barro lee with life expectancy and then merge this data with suicide rate

Drop extra column that are generated by the merge like year, country name(Entity) and country code

```
In [70]:  merged_data = barro_lee_data.merge(life_expectancy_data,left_on=['country','year'],right_on=['Entity','Year'],how = 'left').reset
          merged_data.drop(['Year','Entity','Code'],axis=1,inplace=True)
          merged_data = merged_data.merge(suicide_death_rate_data,left_on=['country','year'],right_on=['Entity','Year'],how = 'left').reset
          merged_data.drop(['Year','Entity','Code'],axis=1,inplace=True)
```

Related Work

Global educational attainment dashboard | The World Bank

illustrating the distribution of educational attainment in over 100 countries from 1970 to 2010. Viewers can select a country, time period, and/or age group below to view the data.

Related Work

Gender ratios for mean years of schooling

Female-to-male ratio of average years of schooling, expressed in percents. All education levels for population aged 15-64. Regional estimates are population-weighted averages.

Eastern Europe
Latin America and the Caribbean
Advanced Economies
Middle East and North Africa
Asia and the Pacific
Sub-Saharan Africa

Source: Lee and Lee (2016)

CC BY

1870   2010   CHART   DATA   SOURCES

Gender ratio in educational attainment | Our World in Data

shows the evolution of female-to-male ratios of educational attainment (mean years of schooling) across different world regions.There has generally been a strong upward trend in the gender ratios across all world regions, which indicates that the inequality between men and women in access to education has been declining.

Related Work

Productivity vs. educational attainment, 1950 to 2010
Figure illustrates educational attainment (average years of total education) versus productivity (GDP per hour worked).
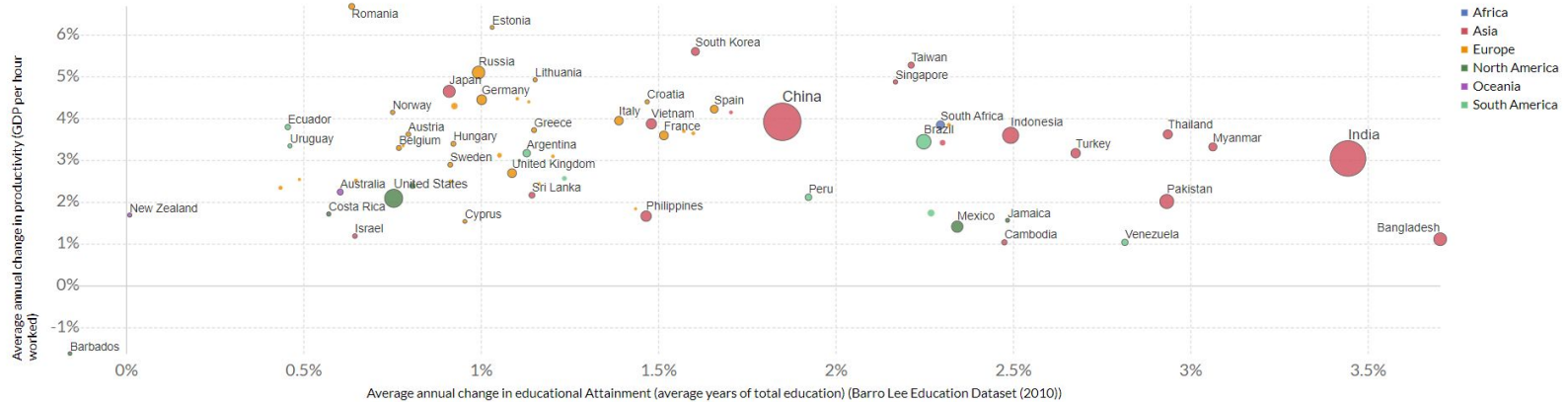
Annual change in productivity vs educational attainment | Our World in Data

This figure illustrates average change in educational attainment (average years of total education) versus productivity (GDP per hour worked) to see whether they have a correlation. Population is encoded in size.

Prototyping Process

Question 1 Does life expectancy correlate with education attainment?

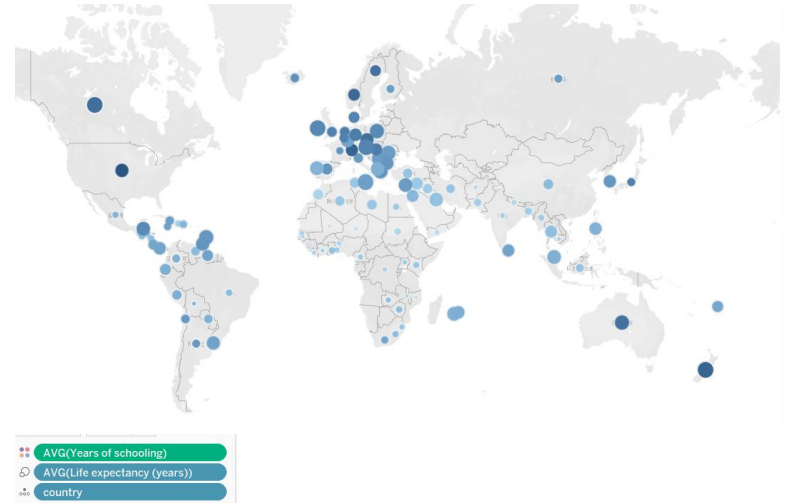1a- Globally
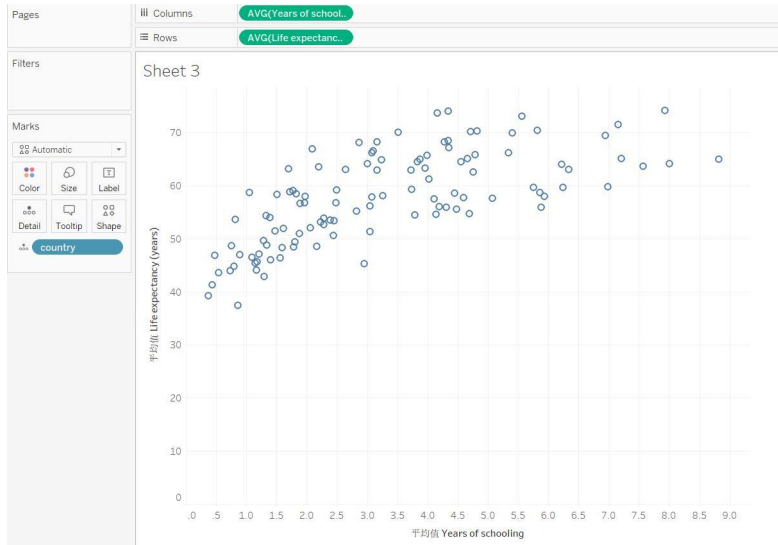
Prototyping Process

Question 1 Does life expectancy correlate with education attainment?

1b- Regionally
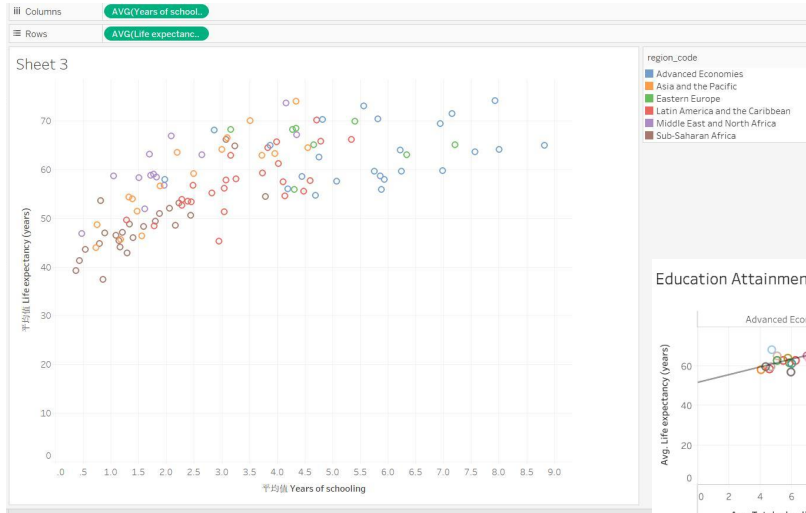
Education attainment in the world

AVG(Years of schooling)
0.375   8.828
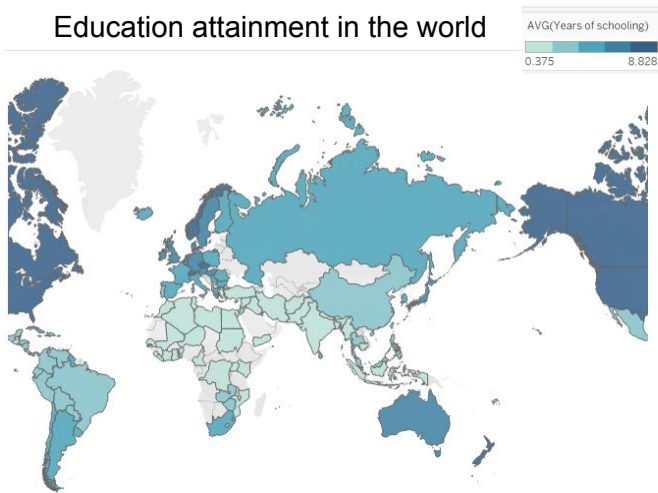
Education attainment and life expectancy in the world

AVG(Years of schooling)
AVG(Life expectancy (years))
country

Prototyping Process

Question 1 Does life expectancy correlate with education attainment?

1b- Regionally

Regions:

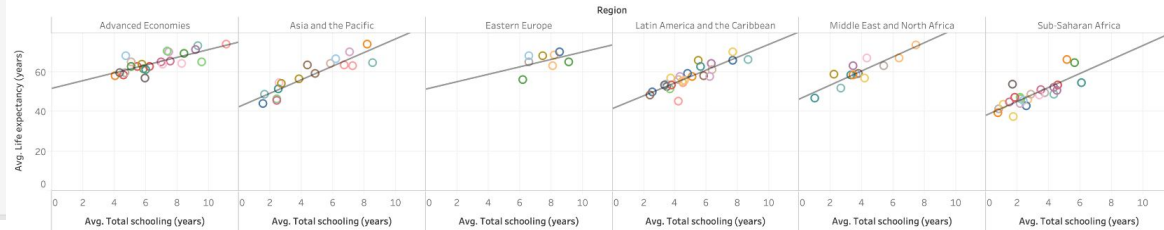Advanced economies
Asia and the Pacific
Eastern Europe
Latin America and the Caribbean
Middle East and North Africa
Sub-Saharan Africa

Prototyping Process

Question 2 Does suicide rate correlate with education attainment?

2a- Globally

Prototyping Process

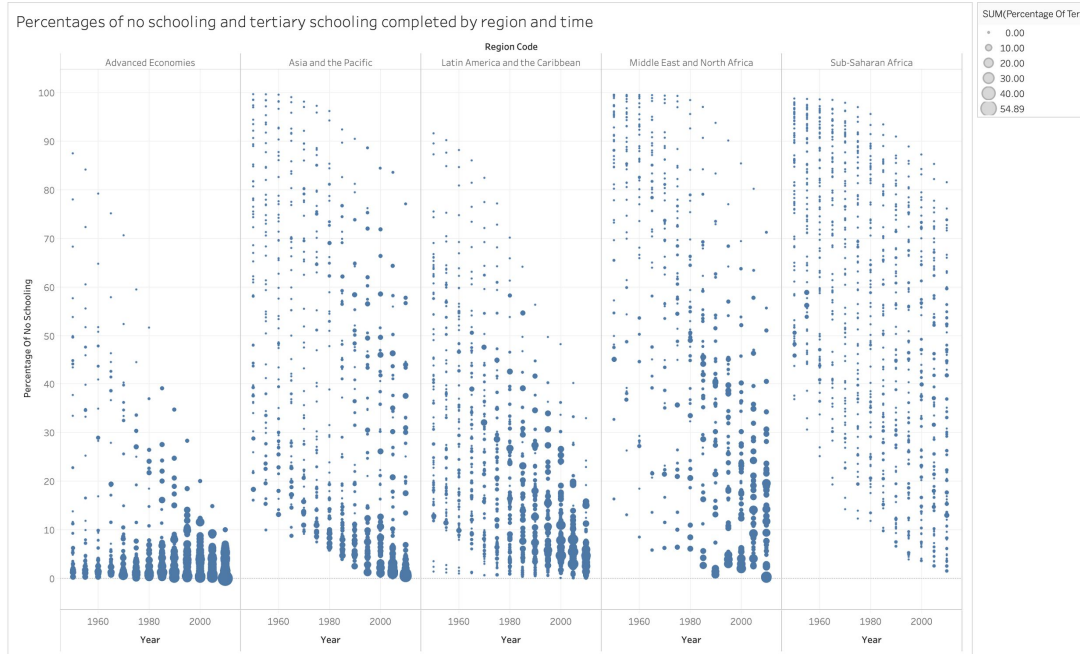**Question 2** Does suicide rate correlate with education attainment?

2b - Regionally

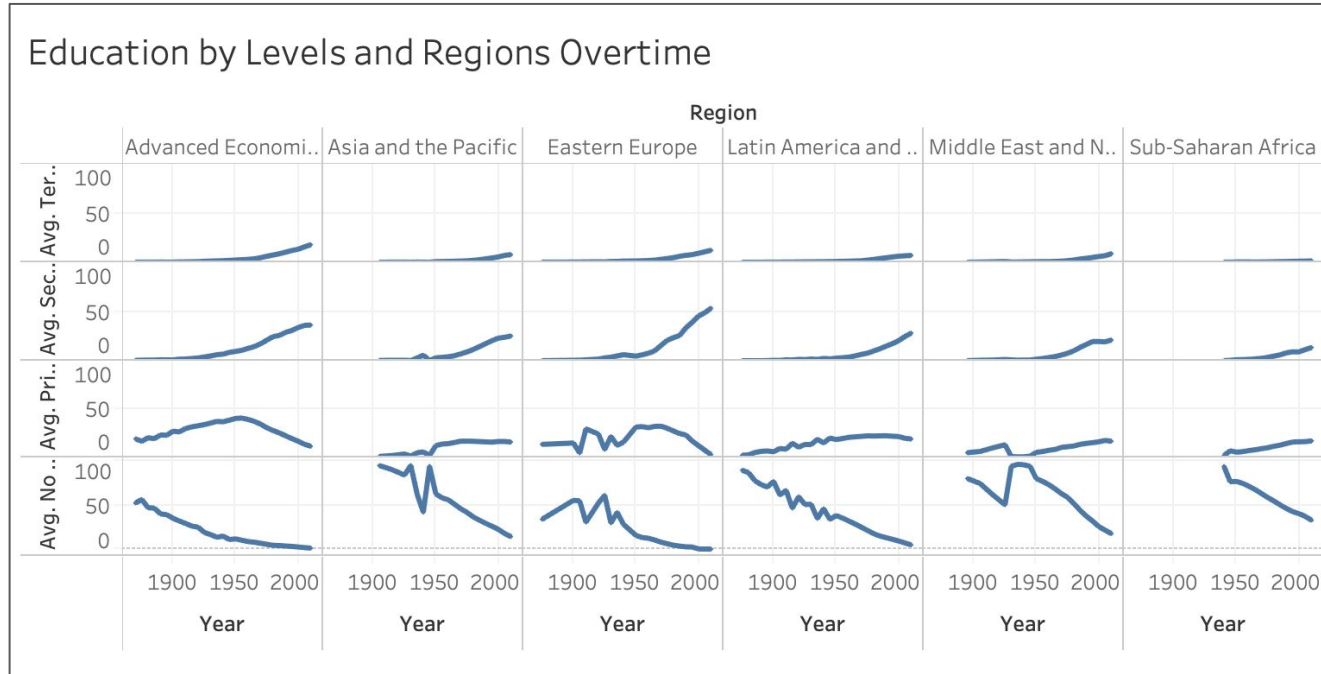Education Attainment versus Suicide Rate
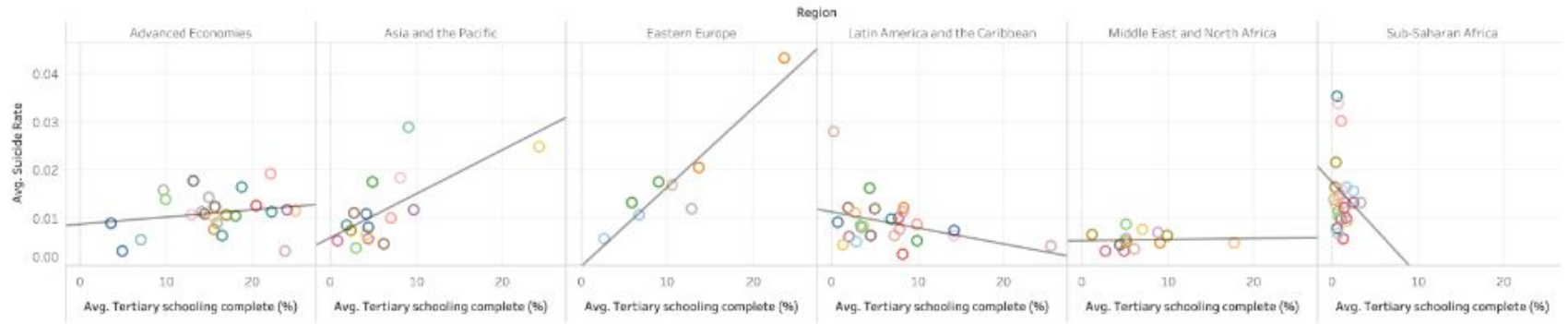
Prototyping Process

Question 3 How do "years of no schooling" and "percentages of tertiary complete" change over time in different countries?
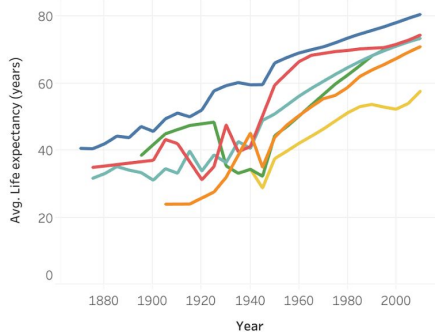

Percentages of no schooling and tertiary schooling completed by region and time

Prototyping Process

Redefined Question 3 How different levels of education change over time in different regions?

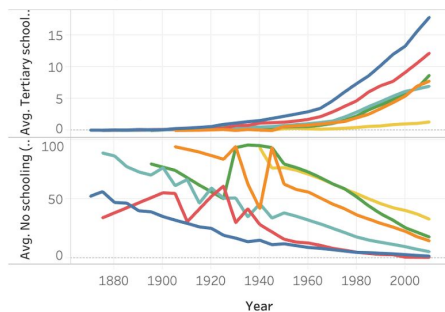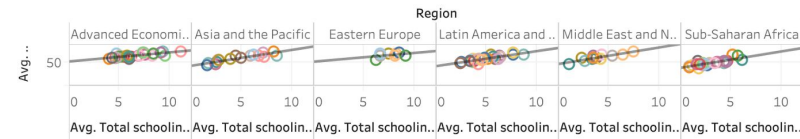Education by Levels and Regions Overtime

Draft Dashboard

Usability Testing

Target User Profile:

- **Name:** Jackie

- **Gender:** Male

- **Age:** 25

- **Occupation:** PhD student in Economics

- **Exposure to data visualization:** Intermediate-high level of proficiency in data visualization tools; high familiarity with different types of data visualizations

- **Knowledge of Education Development:** Basic knowledge of historical trend; general interest in learning more about it

Usability Testing

Discoveries

- Education attainment overview chart is disorienting with too many small sections jumbled together; not sure about what to make of the tertiary education attainment & suicide rate chart

- Lack of interrelatedness between charts on the same dashboard

- More interactivity is needed -> filter / drop-down menus for users to curate the view to their own liking

" What kind of story do we want to tell? "

Discussion [Progress so far and what is left]

- Refine questions and story
- Prototyping -> Dashboard
- More usability tests
- Need to refine our prototype based on the feedback from the usability test
- Visual exploration of data

Thank you!

Appendix

Insights so far

- Correlation between avg life expectancy and education attainment (avg total yrs of schooling) is +ve for all other regions <Susan>
- Trends are also sometimes driven by outliers (e.g. Czech Republic's change in avg total yrs of schooling driving that of the whole eastern europe region) <Susan>
- Illiteracy is reducing in Sub-Saharan Africa, however the reduction is due to primary education and not tertiary education <Anqi>
- There is a positive correlation between life expectancy and education attainment <Shiyao>
- Life expectancy has a hockey stick pattern for mostly sub-saharan africa, where it decreases in 1990s and picks up in 2000s <Amitabh>
- Suicide rate and education attainment correlation by country vs years for Eastern Europe (as-is) <Anqi>
- From 1990 to 2010, advanced economies experience the most sharp increase in percentage of tertiary (literacy), sub saharan area almost has no increase in literacy, with suicide rate increased a lot.

Things to try/explore

- Distribution of years of education by region
- There seem to be two/three lines that can explain correlation between education and suicide rate. Explore more to discover any correlation pattern
- Any correlation with population - life expectancy or education attainment
- Suicide rate and education attainment correlation by country vs years (which is statistically right to compute?)
- Dig into cases of outlier countries -> e.g. South Africa(in the past 10 years), Czech Republic (Driving the average total years of schooling of Eastern Europe, rate exceeds that of Advanced Economies)
- D3 as an option
- Anqi -> schedule usability testing sessions on existing charts (ideally integrated into a dashboard)