

Final Project

Mac Owusu & Divakar Borra

4/28/2022

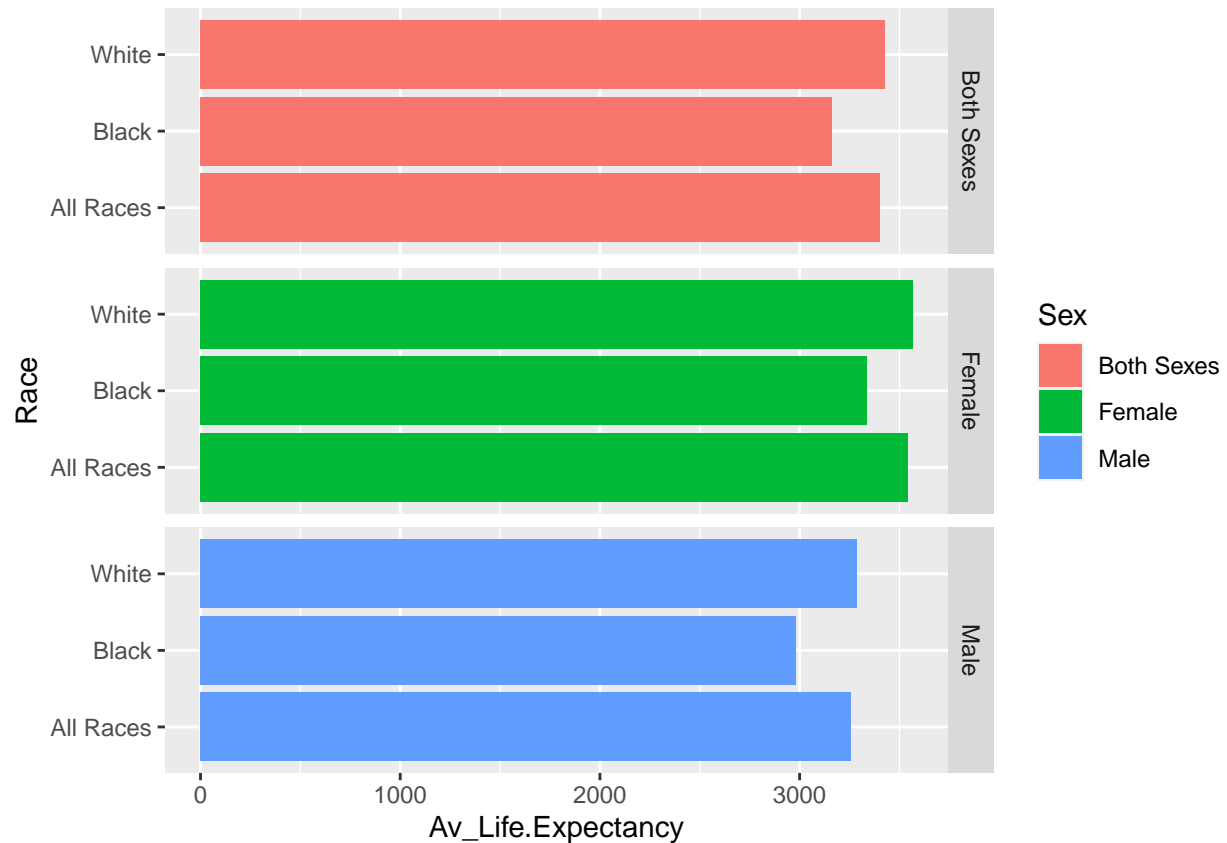
1. Predicting the Scope of Life Expectancy using Past Life Expectancy Data
2. Divakar Borra & Mac Owusu
3. April 29, 2022
4. Life expectancy data from now to predict how life expectancy may increase or decrease in the near future, we'll do this by looking at diseases and other factors that affect the length of an individuals life span.
5. Change to life expectancy because life expectancy becoming prominent due to disease EX. covid, interested in making a prediction on how life expectancy will increase or decrease (covid). We decided to explore the two datasets by using visualizations grpahs such as barplots, line graphs, and demsity curves. We wanted to see the association with certain variables to judge on what models we want to use for late in our project.
6. Finding a dataset that matched our interest, was large enough to fit, whether if it was tidy or not comparing the US life expectancy to the rest of the world using numerous variables to predict whether life expectancy will increase or decrease in the world
7. Graphs and explanations below
8. <https://www.kaggle.com/datasets/nairaminasyan/nchs-death-rates-and-life-expectancy-at-birth> <https://www.kaggle.com/datasets/just249/lifeexpectancydatacsv>

```
library(ggplot2)
```

```
WRLD_lifeExp <- read.csv("~/Info248/Final Project/Life Expectancy Data.csv")  
US_lifeDeathExp <- read.csv("~/Info248/Final Project/Death_rates_and_life_expectancy_at_birth.csv")
```

```
ggplot(US_lifeDeathExp, aes(x=Av_Life.Expectancy,Race,fill=Sex))+geom_bar(stat="identity")+  
facet_grid(vars(rows=Sex))
```

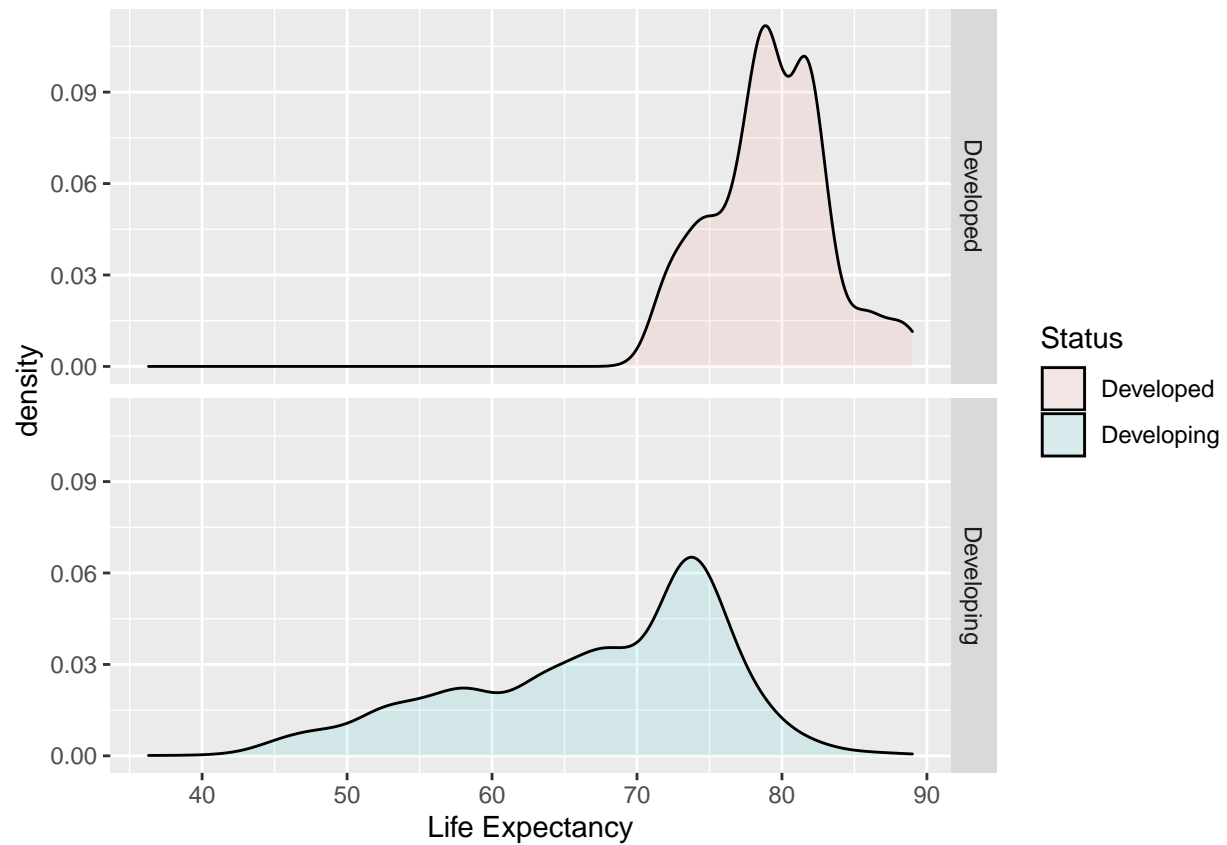
```
## Warning: Removed 9 rows containing missing values (position_stack).
```



> First Graph In this graph we decided to visualize the life expectancy across the world, we wanted to see how Life Expectancy is affected by an individuals sex and race using a bar plot which was faceted by sex.

```
ggplot(WRLD_lifeExp, aes(x=Life_expect, fill=Status))+
  geom_density(alpha=0.1)+
  facet_grid(vars(rows=Status))+
  xlab("Life Expectancy")
```

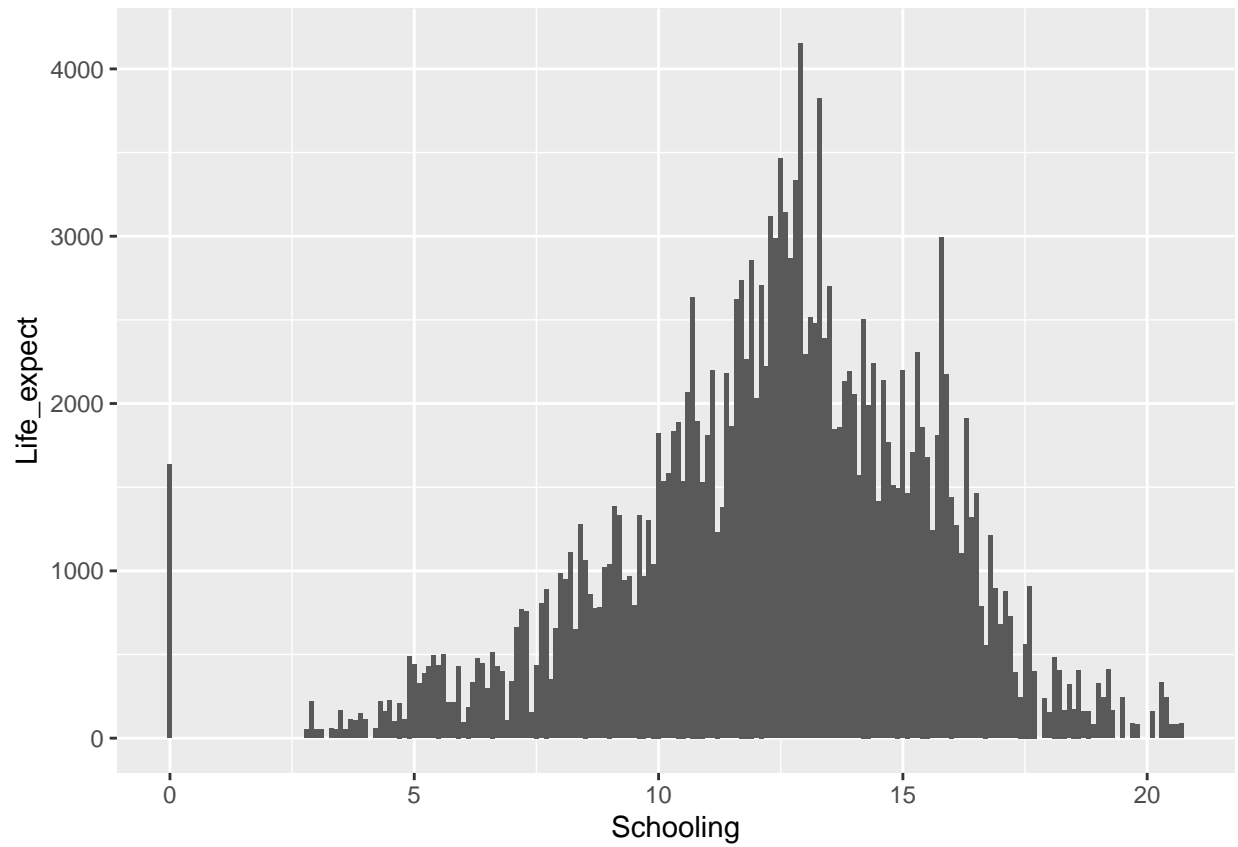
Warning: Removed 10 rows containing non-finite values (stat_density).



> Second Graph The second graph pulls data from the WRLD_lifeExp dataset that shows the probability of life expectancy in a developed and developing country. This is facted by status.

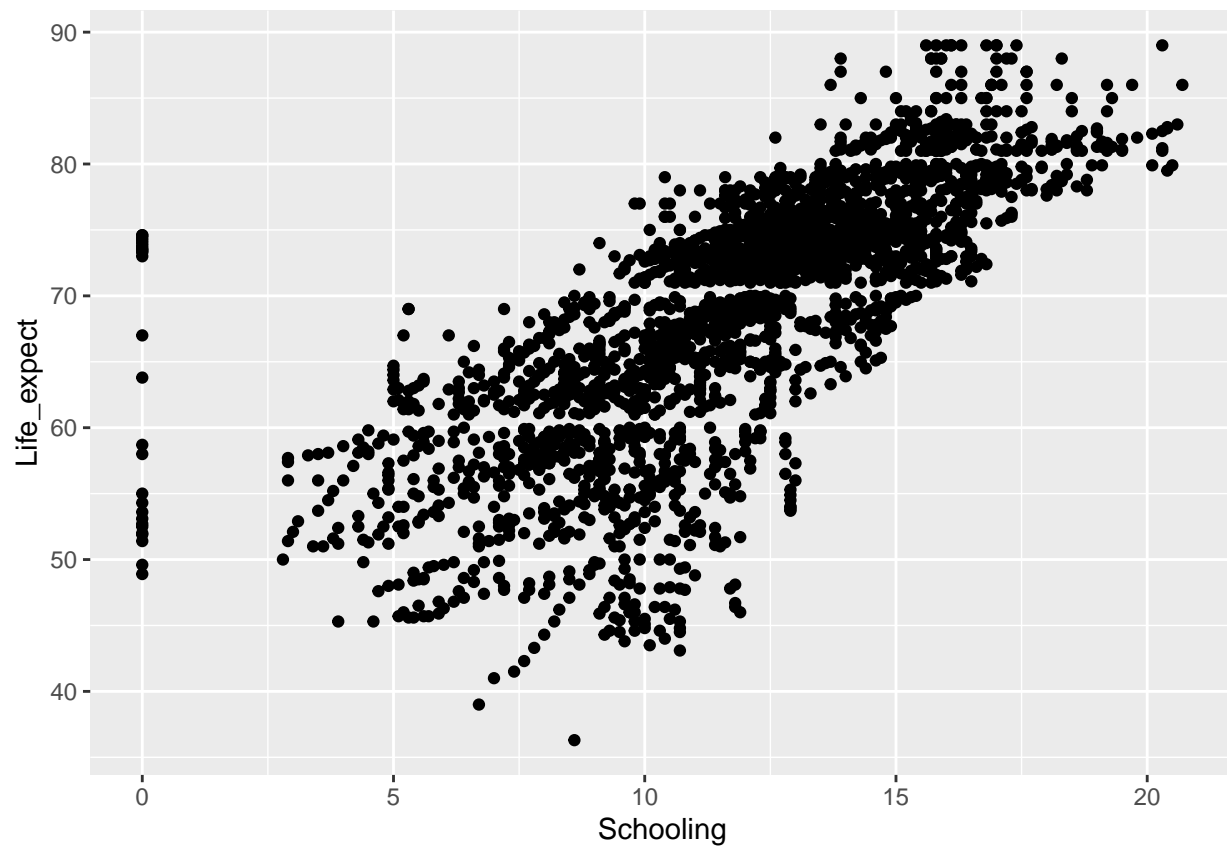
```
ggplot(WRLD_lifeExp, aes(x=Schooling,Life_expect))+
  geom_bar(stat="identity")
```

```
## Warning: Removed 170 rows containing missing values (position_stack).
```



```
ggplot(WRLD_lifeExp, aes(x=Schooling,Life_expect))+  
  geom_point()
```

```
## Warning: Removed 170 rows containing missing values (geom_point).
```



Third & Fourth Graphs (Same Images Different Graphs) The last two graphs look at the data where schooling has an impact on life expectancy. We wanted to see how schooling and education impact life expectancy in certain countries.