DSC680 - Applied Data Science

24.08.10 Week Ten

Milestone Three

Exploring Ways to Curb the Suicide Rate in the Republic of Korea

Professor Iranitalab

Ross Kim-Schreck

```
# 08.02.01.01
# read csv
# assign variable
# dt01

dt01_death_cause_gend________00 = pd.read_csv('101_DT_1B34E09_20240718123555.csv')
```

```
# 08.02.01.02
# read csv
# assign variable
# dt02

dt02_death_cause_geo_________00 = pd.read_csv('101_DT_1B34E12_20240718123805.csv')
```

```
# 08.02.01.03
# read csv
# assign variable
# dt03

dt03_who_suicide_____________00 = pd.read_csv('who_suicide_statistics.csv')
```

```
# 10.02.52.13
# read csv
# assign variable
# dt35

dt35_suic_reason_____________00 = pd.read_csv('Reason_and_Attempt_to_Think_Suicide_by_General_Feature_of_older_persons_Over_65_Years_Old__20240801134427.csv', encoding='unicode_escape')
```

```
# 10.02.52.14
# read csv
# assign variable
# dt36

dt36_depr_symptom____________00 = pd.read_csv('Symptom_of_Depression_by_General_Feature_of_older_persons_Over_65_Years_Old__20240801135254.csv', encoding='unicode_escape')
```

```
# 10.02.52.15
# read csv
# assign variable
# dt37

dt37_happiness_world_________00 = pd.read_csv('World Happiness Report.csv')
```

```
# 09.02.02.01
# confirm column names
# dt01

dt01_death_cause_gend_____00.columns
```

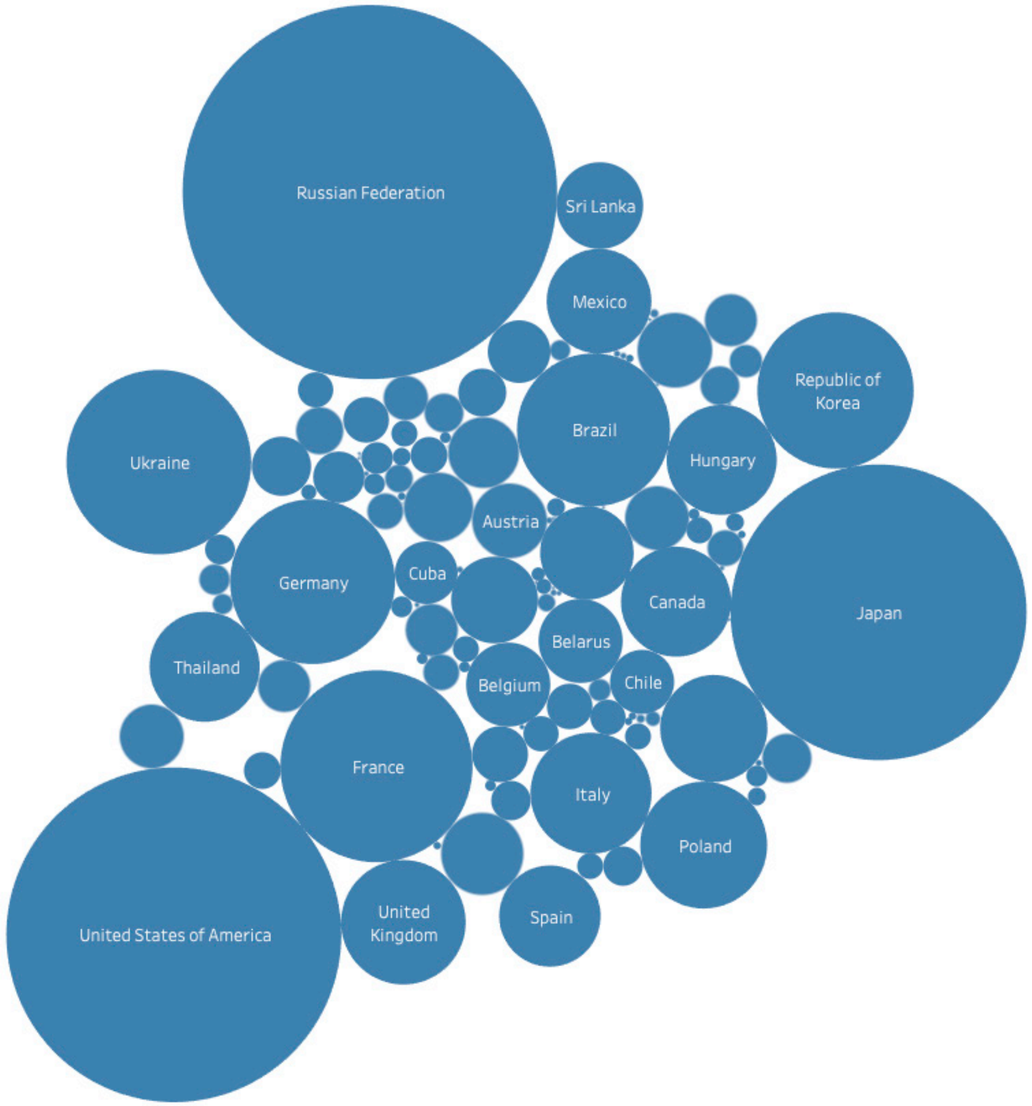
```
Index(['By the cause of death(104 items)', 'By gender', 'By province', 'Item',
       'UNIT', '1983 Year', '1984 Year', '1985 Year', '1986 Year', '1987 Year',
       '1988 Year', '1989 Year', '1990 Year', '1991 Year', '1992 Year',
       '1993 Year', '1994 Year', '1995 Year', '1996 Year', '1997 Year',
       '1998 Year', '1999 Year', '2000 Year', '2001 Year', '2002 Year',
       '2003 Year', '2004 Year', '2005 Year', '2006 Year', '2007 Year',
       '2008 Year', '2009 Year', '2010 Year', '2011 Year', '2012 Year',
       '2013 Year', '2014 Year', '2015 Year', '2016 Year', '2017 Year',
       '2018 Year', '2019 Year', '2020 Year', '2021 Year', '2022 Year',
       'Unnamed: 45'],
      dtype='object')
```

```
# 09.02.02.02
# confirm column names
# dt02

dt02_death_cause_geo_____00.columns
```

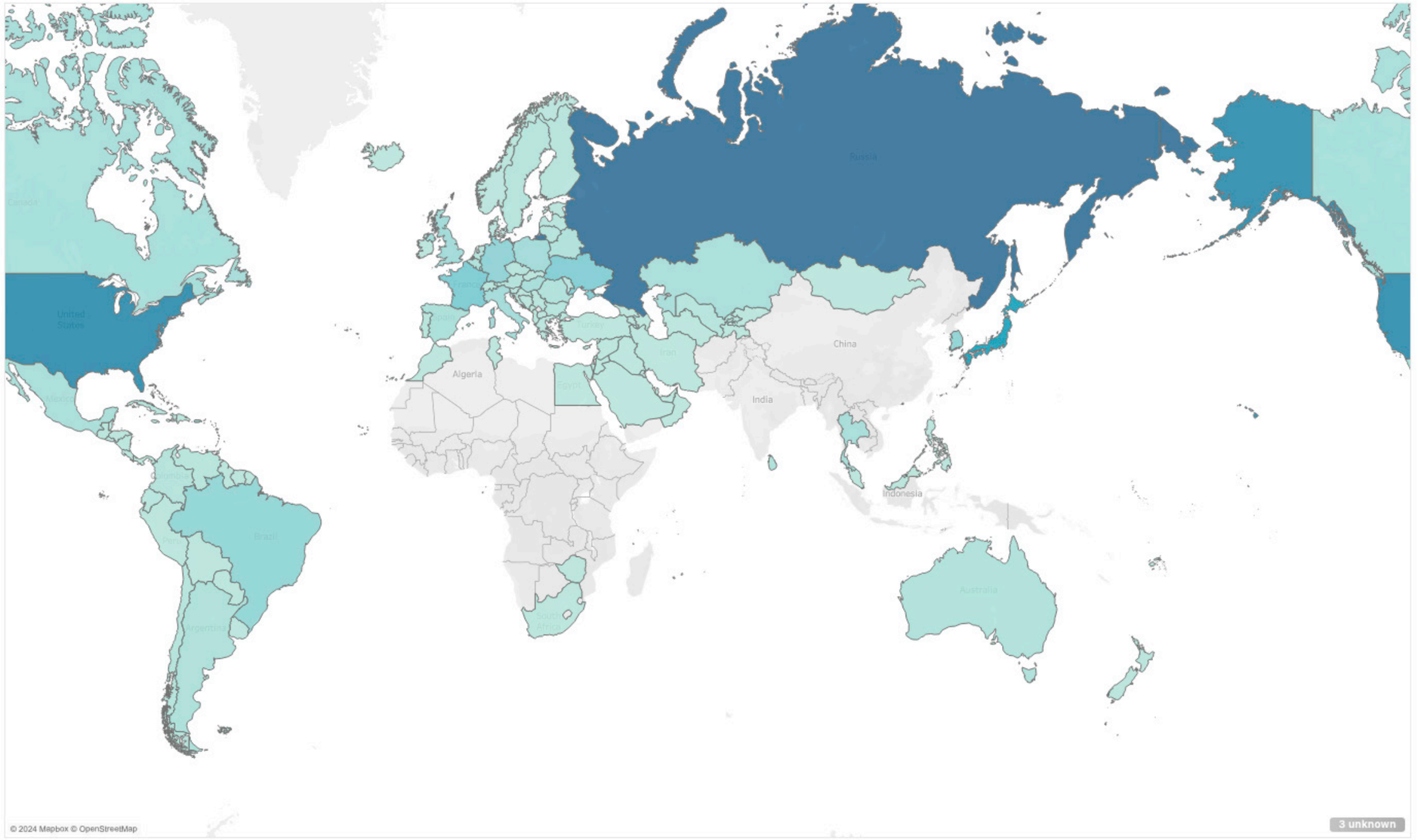
```
Index(['By the cause of death(104 items)', 'By province', 'By gender', 'Item',
       'UNIT', '1996 Year', '1997 Year', '1998 Year', '1999 Year', '2000 Year',
       '2001 Year', '2002 Year', '2003 Year', '2004 Year', '2005 Year',
       '2006 Year', '2007 Year', '2008 Year', '2009 Year', '2010 Year',
       '2011 Year', '2012 Year', '2013 Year', '2014 Year', '2015 Year',
       '2016 Year', '2017 Year', '2018 Year', '2019 Year', '2020 Year',
       '2021 Year', '2022 Year', 'Unnamed: 32'],
      dtype='object')
```

```
# 09.02.02.03
# confirm column names
# dt03
```

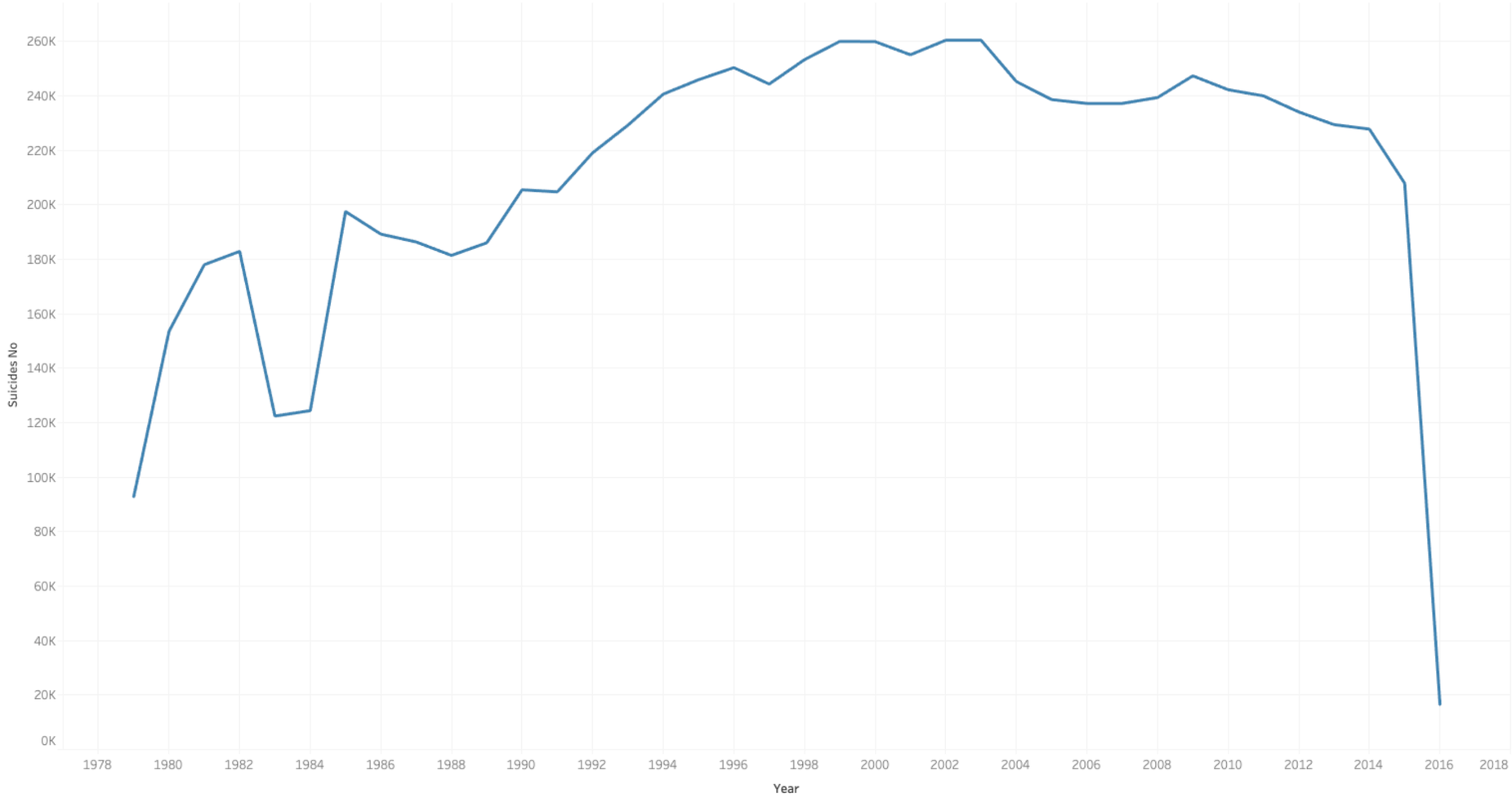
dt03 - Bubble Chart - Number of Suicides



dt03 - Area Map - Number of Suicides



dt03 - Line Graph - Suicides by Year



```
# 10.02.01.01
# read csv
# assign variable
# dt04

dt04_combined_____________00 = pd.read_csv('combined_processed_data.csv')
```

```
# 10.02.01.02
# read csv
# assign variable
# dt05

dt05_stress_general_______00 = pd.read_csv('Degree_of_Stress__General_Life__13_years_old_and_over__20240719092712.csv')
```

```
# 10.02.01.03
# read csv
# assign variable
# dt06

dt06_stress_home__________00 = pd.read_csv('Degree_of_Stress__Home_Life__13_years_old_and_over__20240719092914.csv')
```

```
# 10.02.01.04
# read csv
# assign variable
# dt07

dt07_stress_school________00 = pd.read_csv('Degree_of_Stress__School_Life__13_years_old_and_over__20240719092757.csv', encoding = 'unicode_escape')
```

```
# 10.02.01.05
# read csv
```

```
# 10.02.01.07
# read csv
# assign variable
# dt10

dt10_drinking_20____________00 = pd.read_csv('Drinking__20_years_old_and_over__20240719093241.csv')
```

```
# 10.02.01.08
# read csv
# assign variable
# dt11

dt11_drinking_manage_19_____00 = pd.read_csv('Drinking_and_Health_Management__19_years_old_and_over__20240719093528.csv')
```

```
# 10.02.01.09
# read csv
# assign variable
# dt12

dt12_drinking_manage_20_____00 = pd.read_csv('Drinking_and_Health_Management__20_years_old_and_over__20240719093453.csv')
```

```
# 10.02.01.09
# read csv
# assign variable
# dt13

dt13_suicide_impulse________00 = pd.read_csv('Impulse_to_Commit_Suicide_and_Reasons__13_years_old_and_over__20240719092337.csv')
```

```
# 10.02.01.10
# read csv
# assign variable
# dt14

dt14_suicide_impulse________00 = pd.read_csv('Impulse_to_Commit_Suicide_and_Reasons__13_years_old_and_over__20240719092337.csv')
```

```
# 10.02.01.11
```

```
# 10.02.01.12
# read csv
# assign variable
# dt16

dt16_suicide_reason__________00 = pd.read_csv('Reason_and_Attempt_to_Think_Suicide_by_General_Feature_of_older_persons_Over_65_Years_Old__20240719092517.csv', encoding =
'unicode_escape')
```

```
# 10.02.01.13
# read csv
# assign variable
# dt17

dt17_smoke_drink_19__________00 = pd.read_csv('Smoking_and_Drinking__19_years_old_and_over__20240719093138.csv')
```

```
# 10.02.01.14
# read csv
# assign variable
# dt18

dt18_smoke_drink_20__________00 = pd.read_csv('Smoking_and_Drinking__20_years_old_and_over__20240719093056.csv')
```

```
# 10.02.01.15
# read csv
# assign variable
# dt19

dt19_ph_categories___________00 = pd.read_csv('ph_categories_index.csv')
```

```
# 10.02.01.16
# read csv
# assign variable
# dt20

'''dt20_ph_analysis___________00 = pd.read_excel('ph_Pornhub Analysis year by year.xlsx')'''
```

```
# 10.02.52.01
# read csv
# assign variable
# dt23

dt23_408_03_____________________00 = pd.read_csv('408_DT_40803_N0003_20240801134720.csv')
```

```
# 10.02.52.02
# read csv
# assign variable
# dt24

dt24_408_04_____________________00 = pd.read_csv('408_DT_40803_N0004_20240801134840.csv')
```

```
# 10.02.52.03
# read csv
# assign variable
# dt25

dt25_index_eco_sent_____________00 = pd.read_csv('Economic_Sentiment_Index_20240801135039.csv')
```

```
# 10.02.52.04
# read csv
# assign variable
# dt26

dt26_sadness____________________00 = pd.read_csv('Feeling_sad_or_hopeless_20240801134129.csv')
```

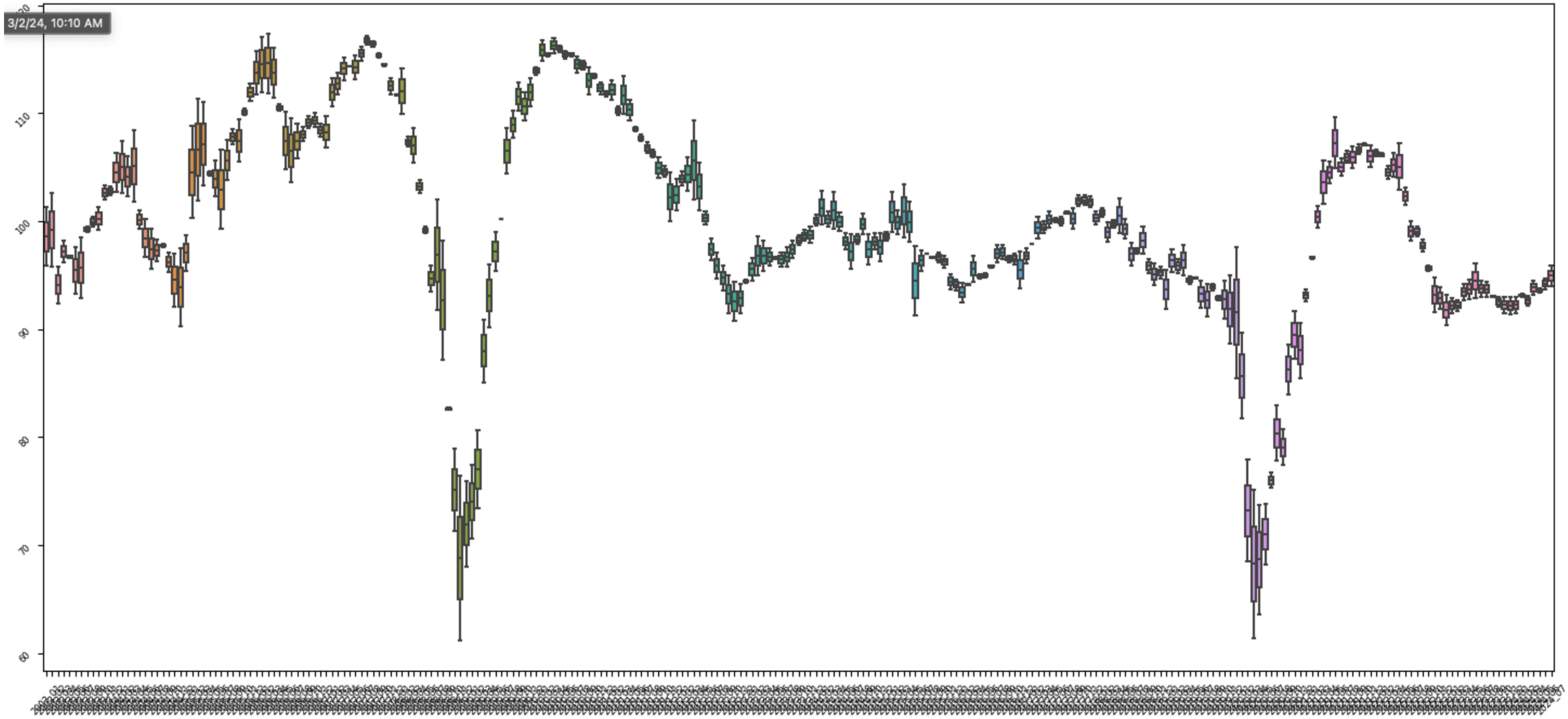
```
# 10.02.52.05
# read csv
# assign variable
```

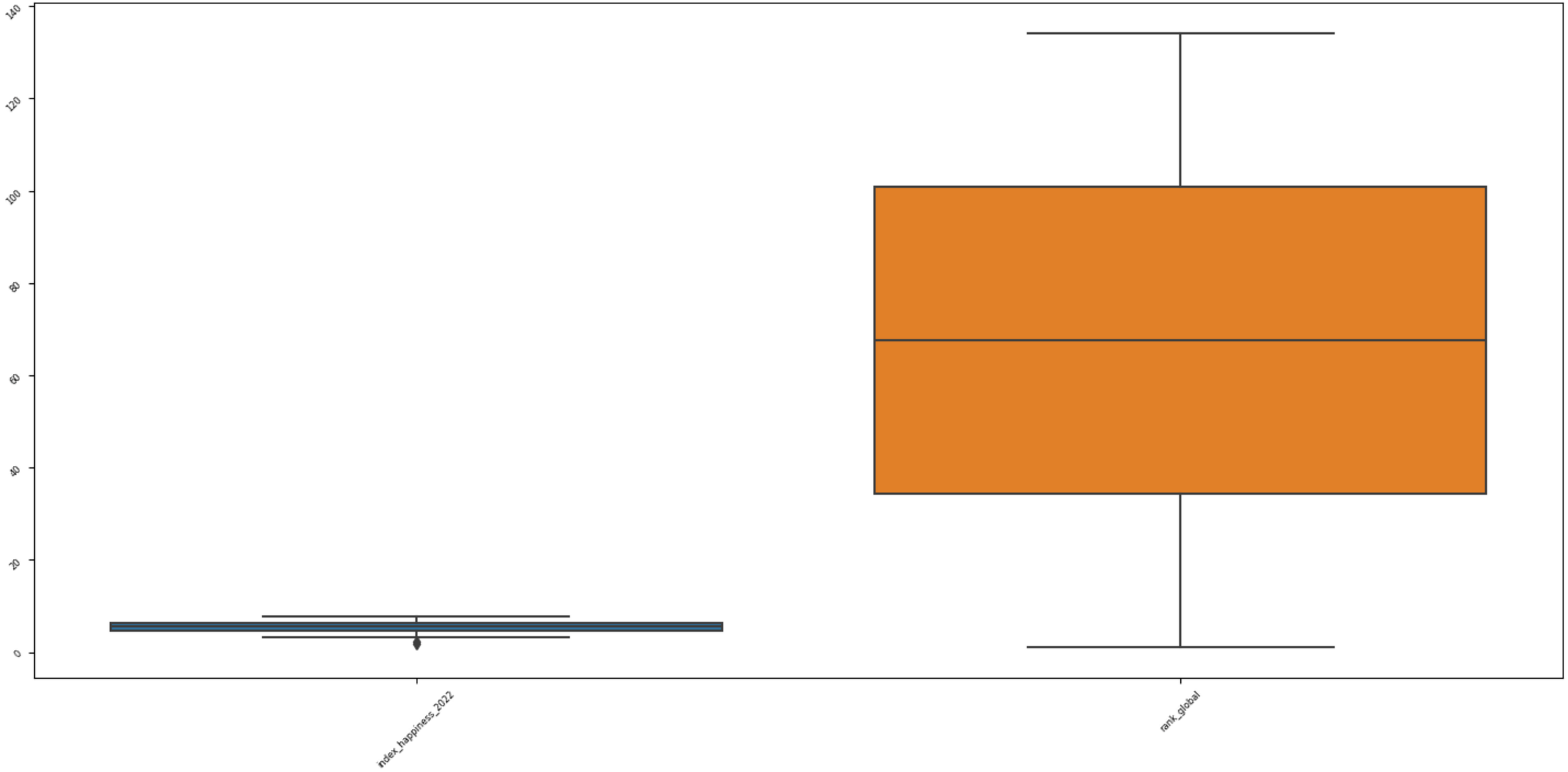
```
# 10.02.52.07
# read csv
# assign variable
# dt29

dt29_happiness_2016__________00 = pd.read_csv('index_happiness_2016.csv')
```

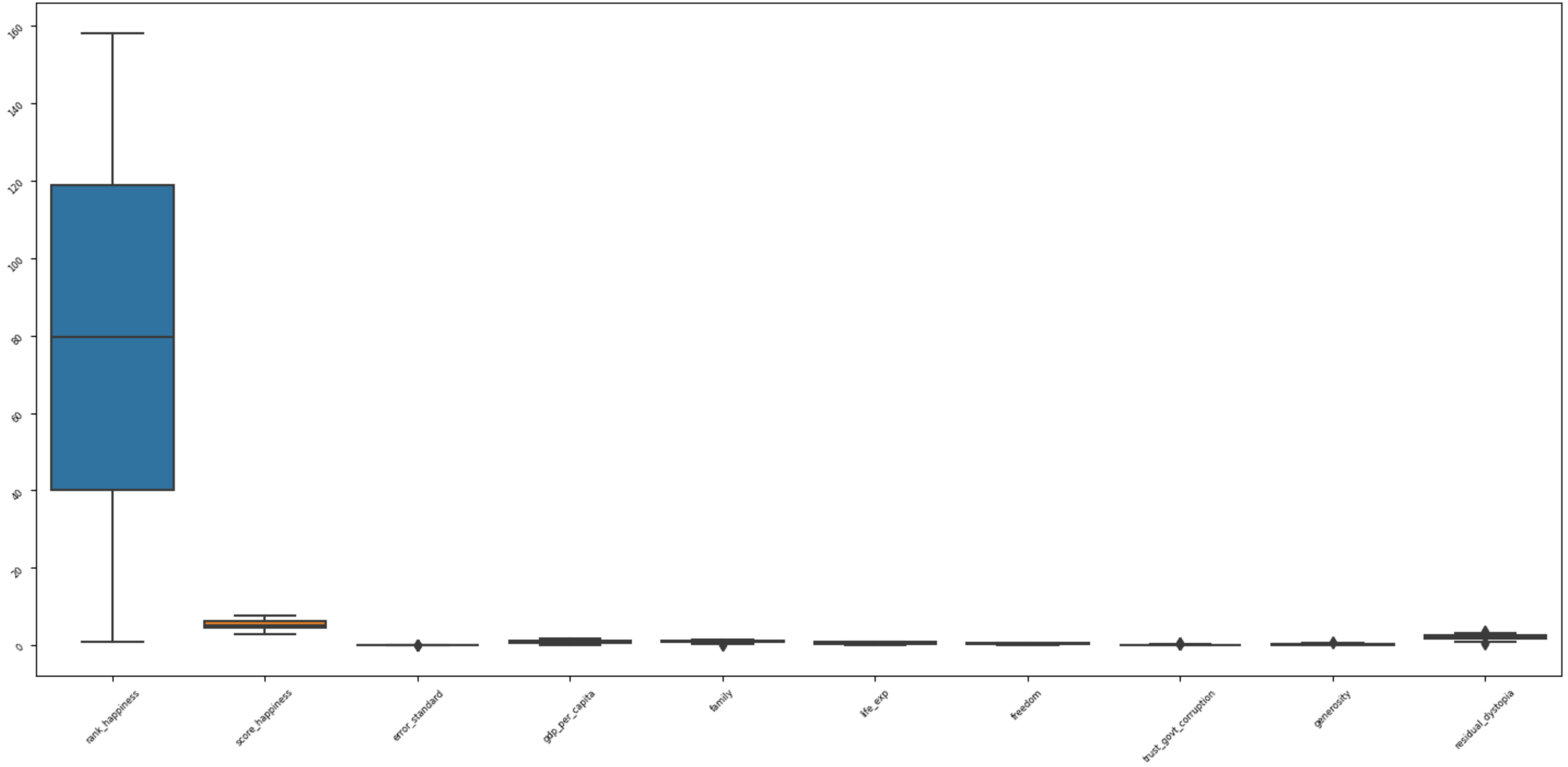
```
# 10.02.52.08
# read csv
# assign variable
# dt30

dt30_happiness_2017__________00 = pd.read_csv('index_happiness_2017.csv')
```

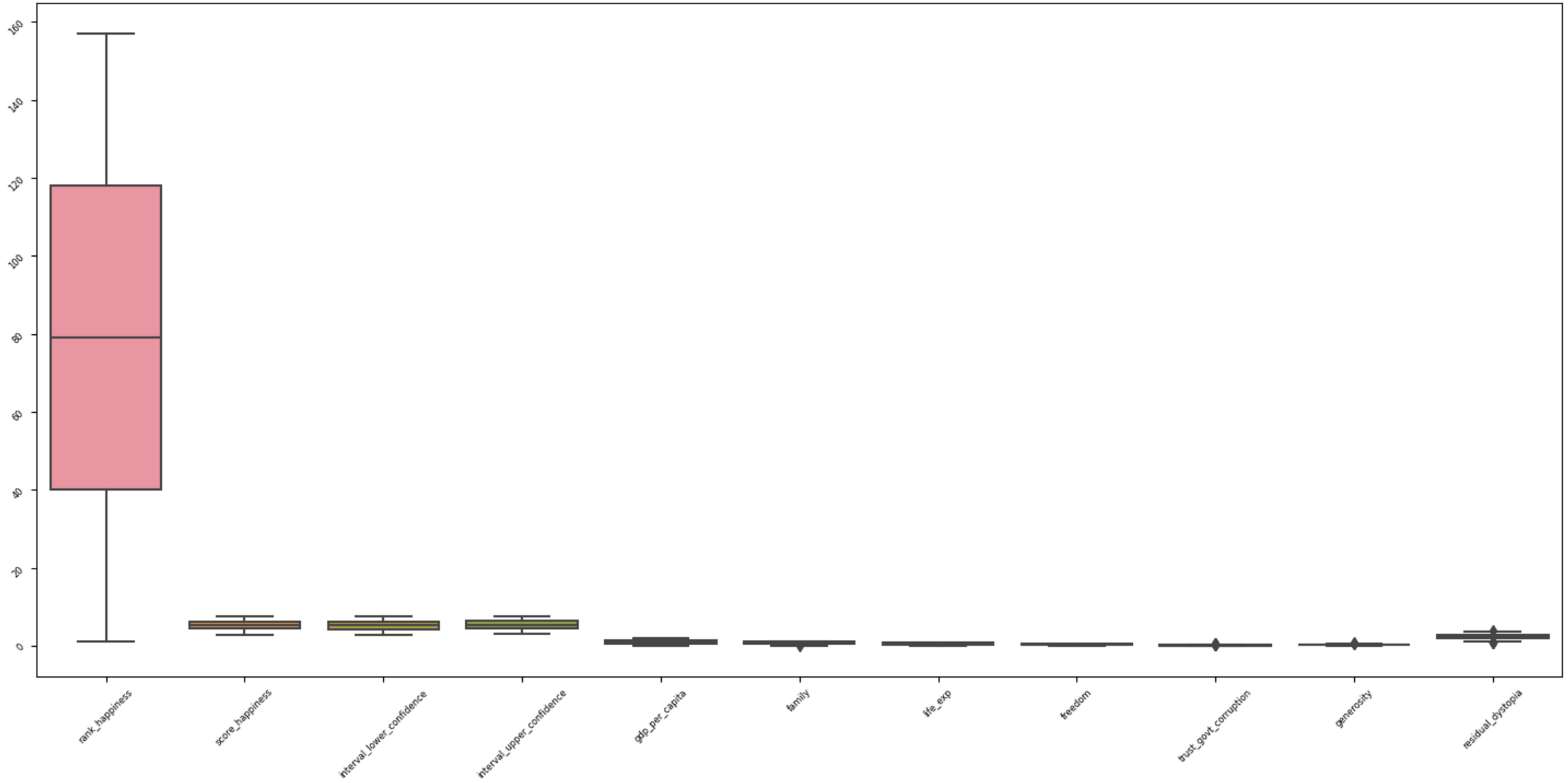
```
# 10.02.52.09
# read csv
# assign variable
# dt31

dt31_happiness_2018__________00 = pd.read_csv('index_happiness_2018.csv')
```

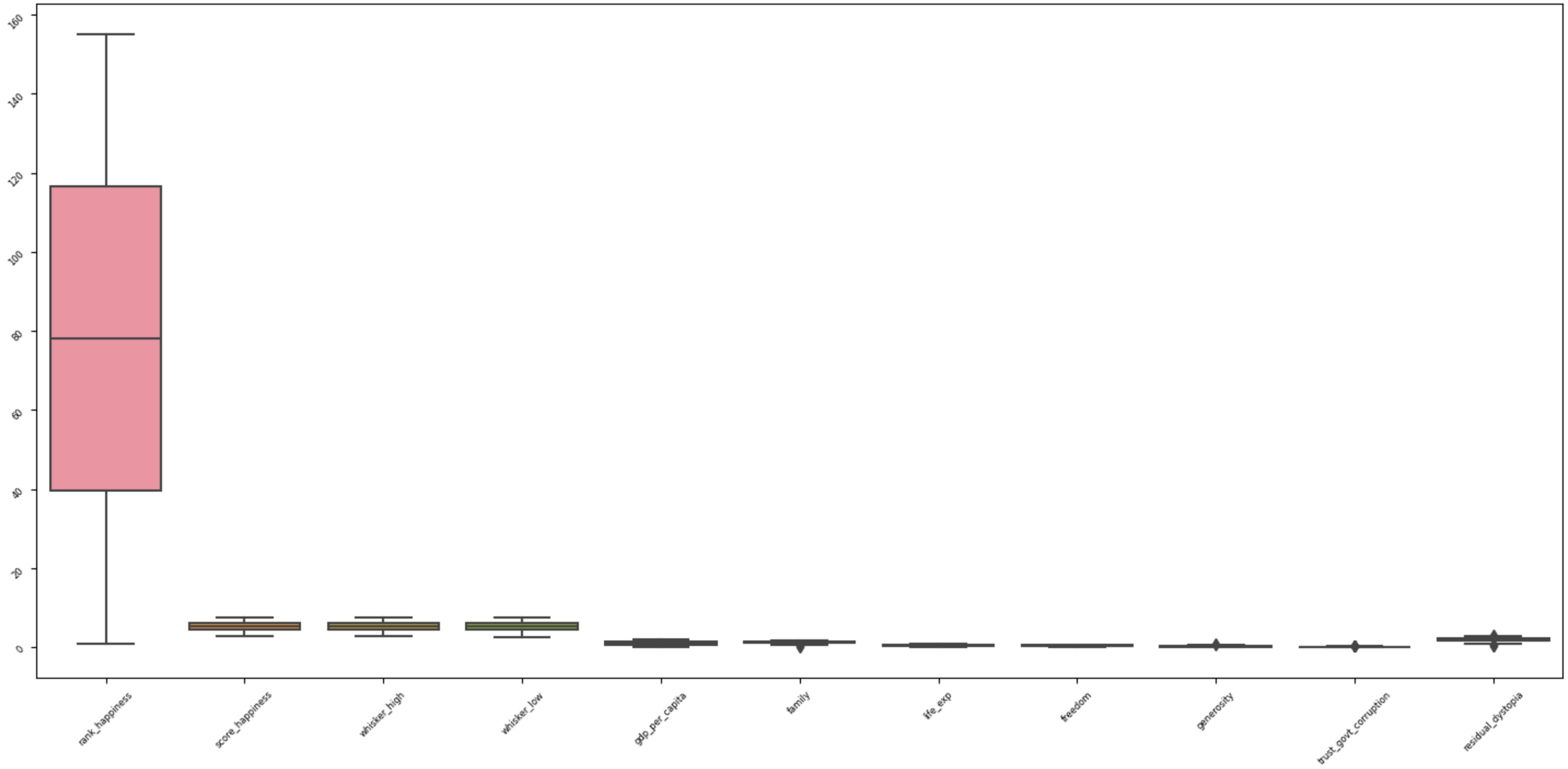
```
# 10.02.52.10
# read csv
# assign variable
# dt32
```

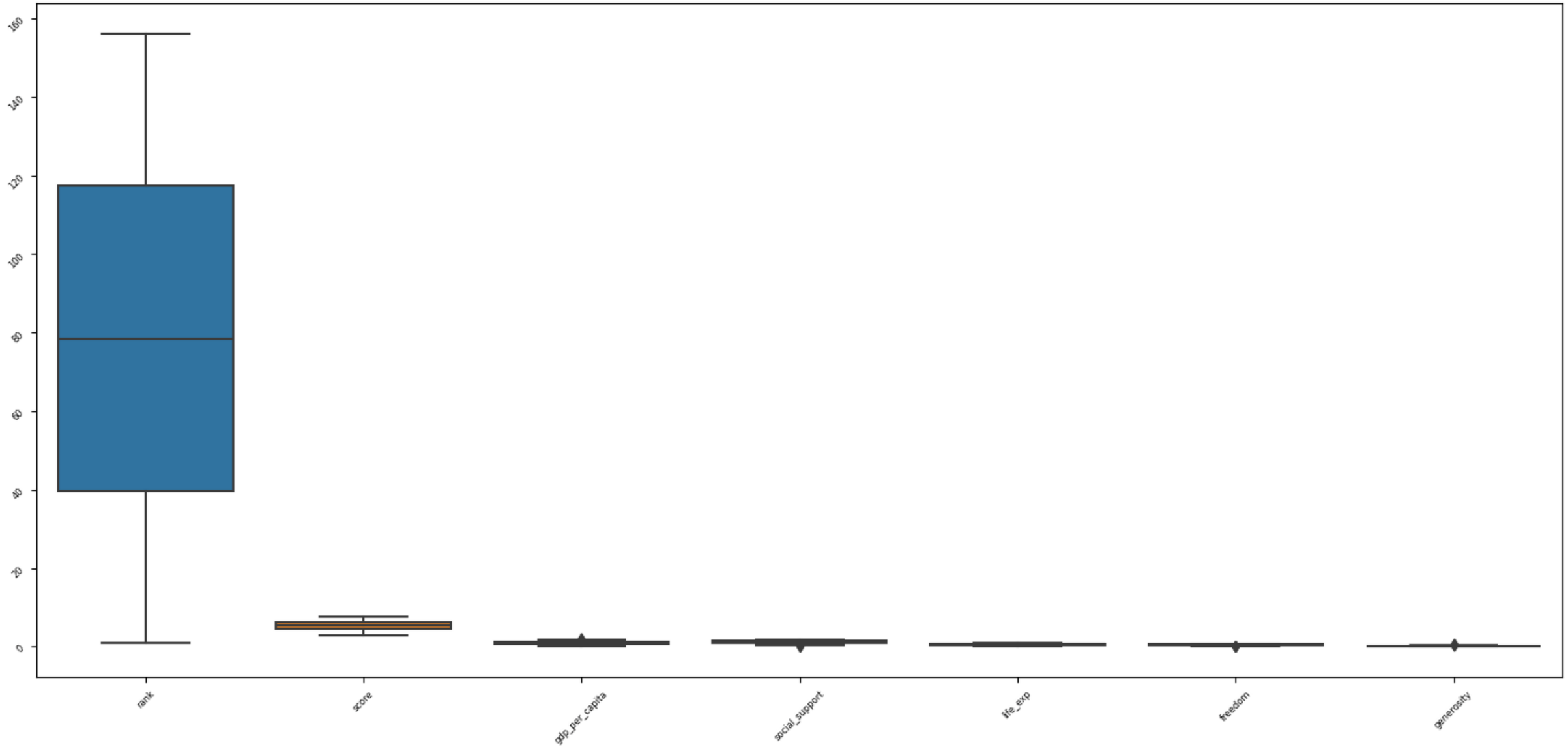

3/2/24, 10:10 AM
110
100
90
80
70
60

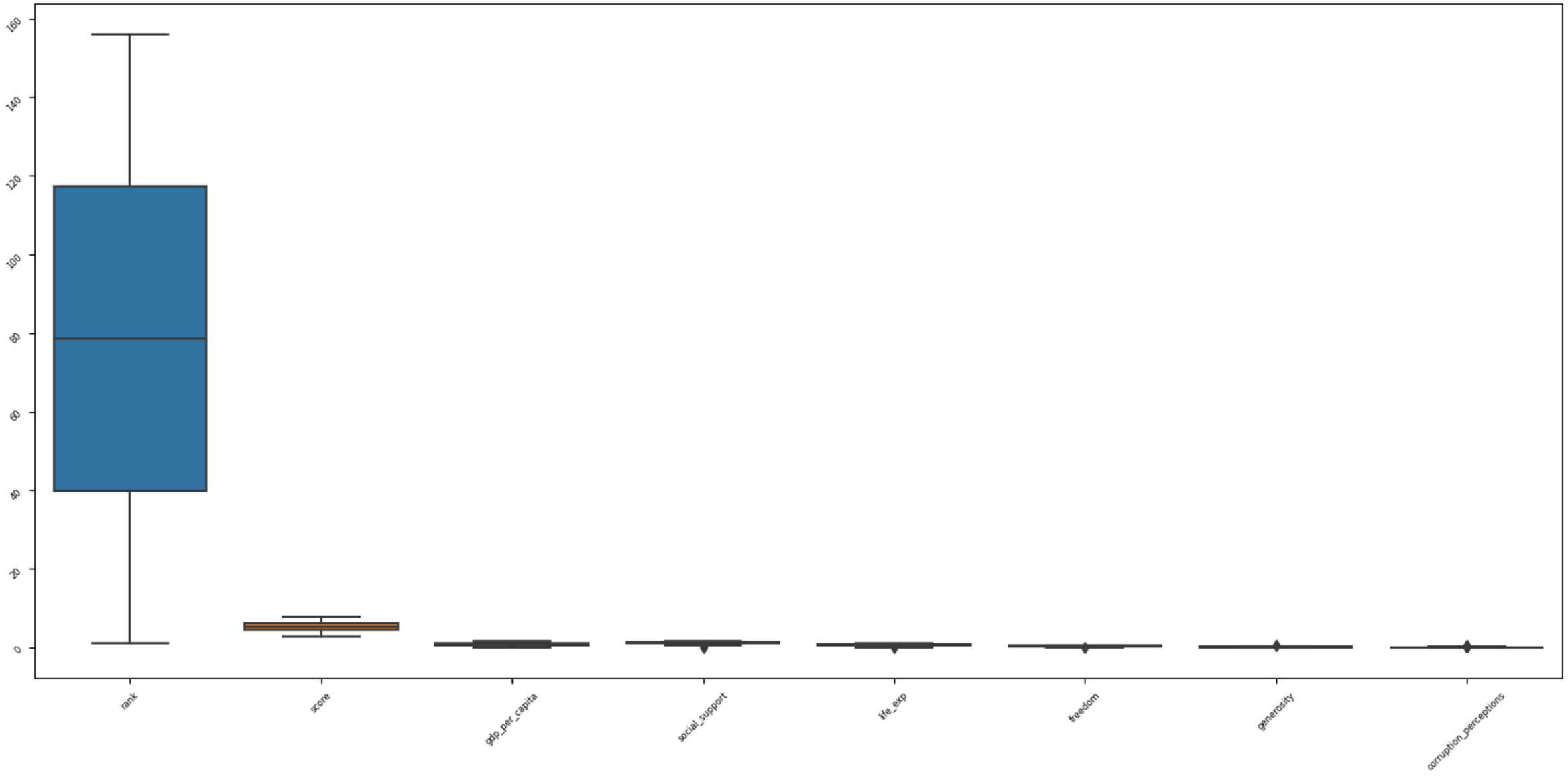

140
120
100
80
60
40
20
0
index_happiness_2022
rank_global


160
140
120
100
80
60
40
20
0
rank_happiness
score_happiness
error_standard
gdp_per_capita
family
life_exp
freedom
trust_govt_corruption
generosity
residual_dystopia

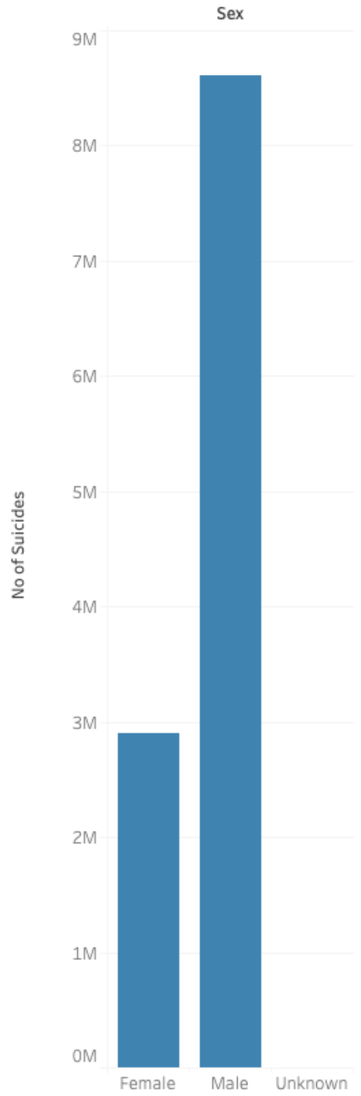

160
140
120
100
80
60
40
20
0
rank_happiness
score_happiness
interval_lower_confidence
interval_upper_confidence
gdp_per_capita
family
life_exp
freedom
trust_govt_corruption
generosity
residual_dystopia


160
140
120
100
80
60
40
20
0
rank_happiness
score_happiness
whisker_high
whisker_low
gdp_per_capita
family
life_exp
freedom
generosity
trust_govt_corruption
residual_dystopia

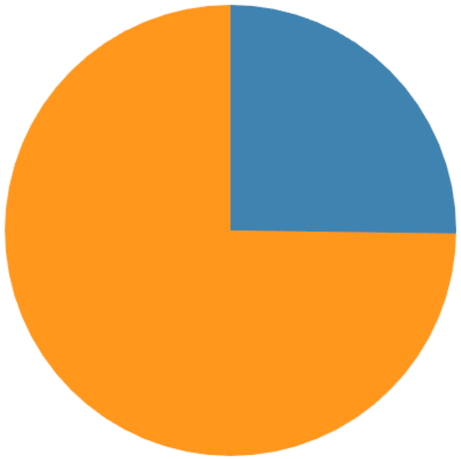

160
140
120
100
80
60
40
20
0
rank
score
gdp_per_capita
social_support
life_exp
freedom
generosity

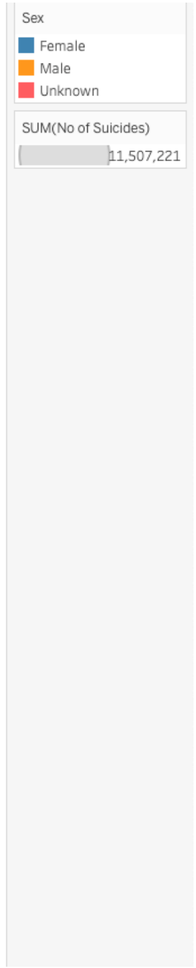

160
140
120
100
80
60
40
20
0
rank
score
gdp_per_capita
social_support
life_exp
freedom
generosity
corruption_perceptions

dt04 - Bar Chart - Suicides by Gender

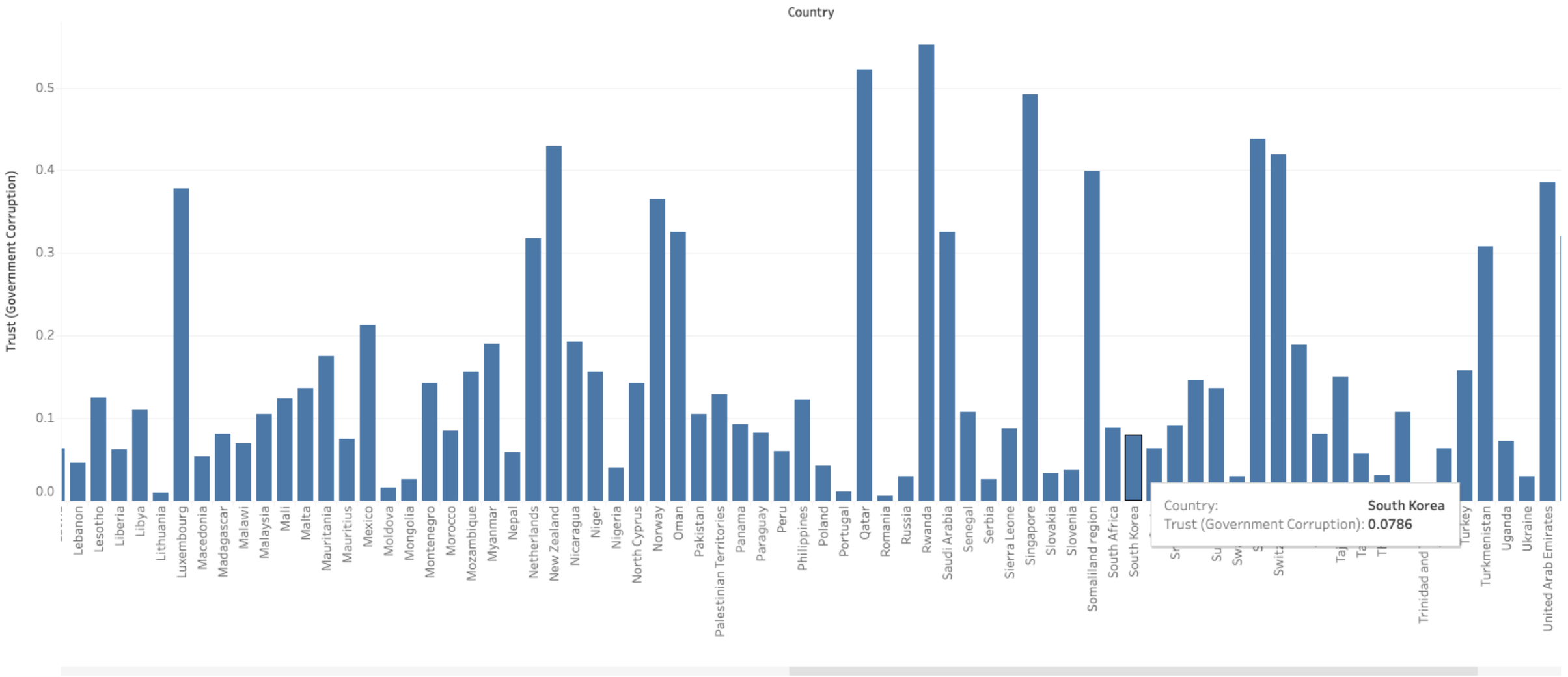


dt04 - Pie Chart - Suicides by Gender

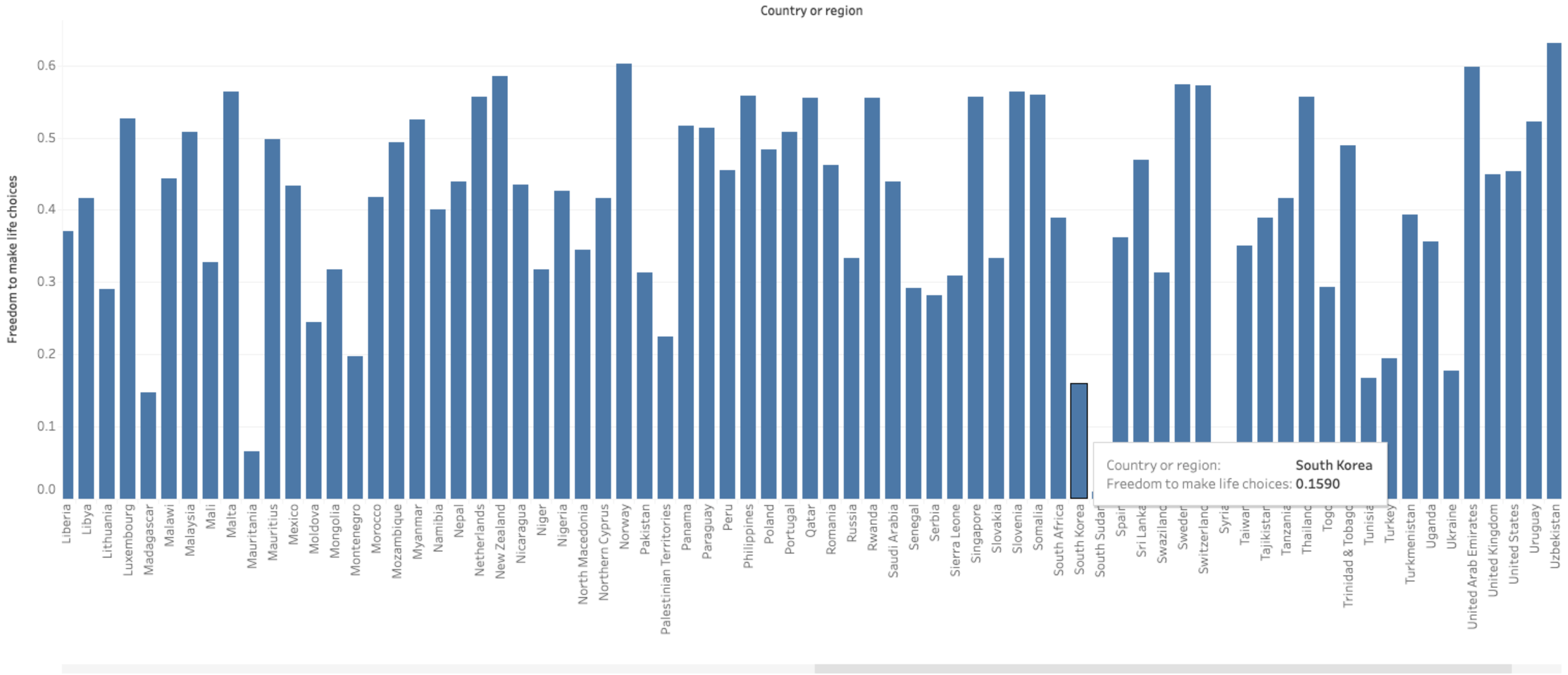




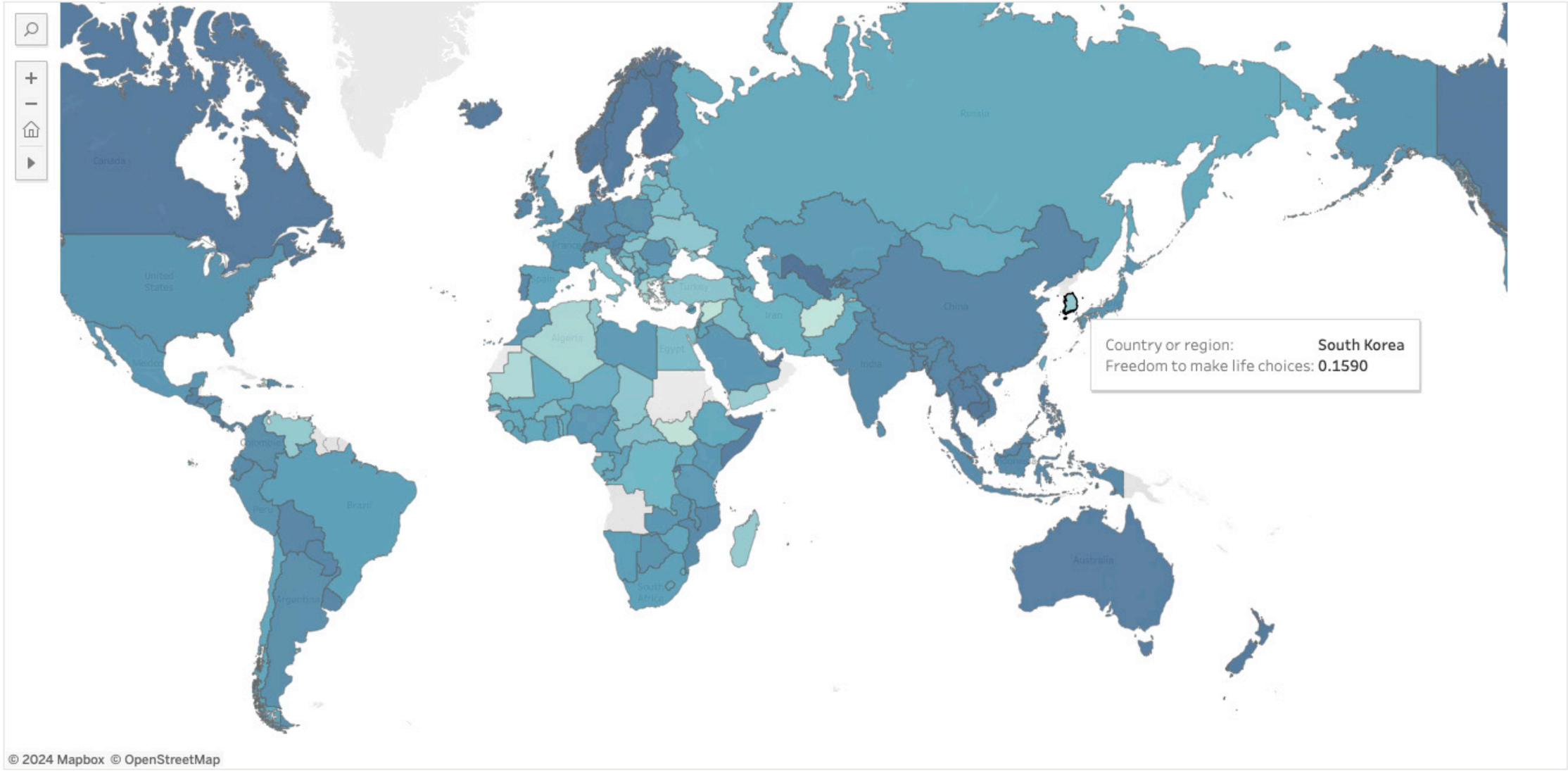
dt28 - Bar Chart 04 Gov't Trust



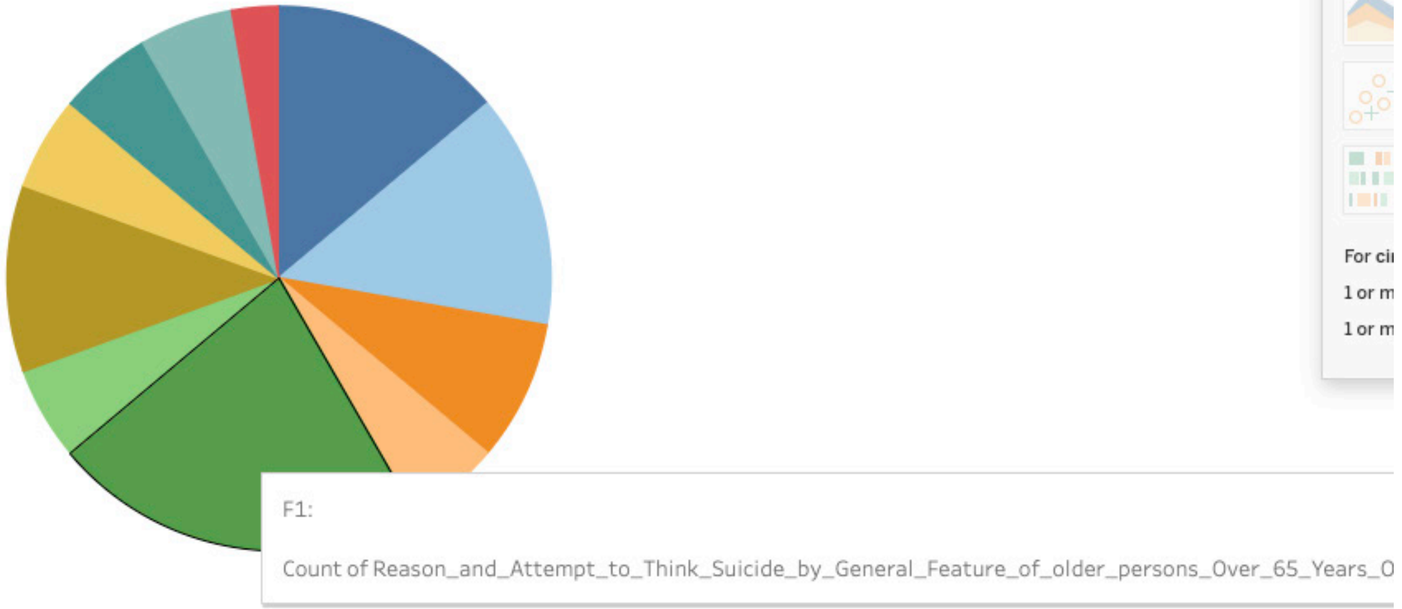
dt32 - Bar Chart 01 Freedom



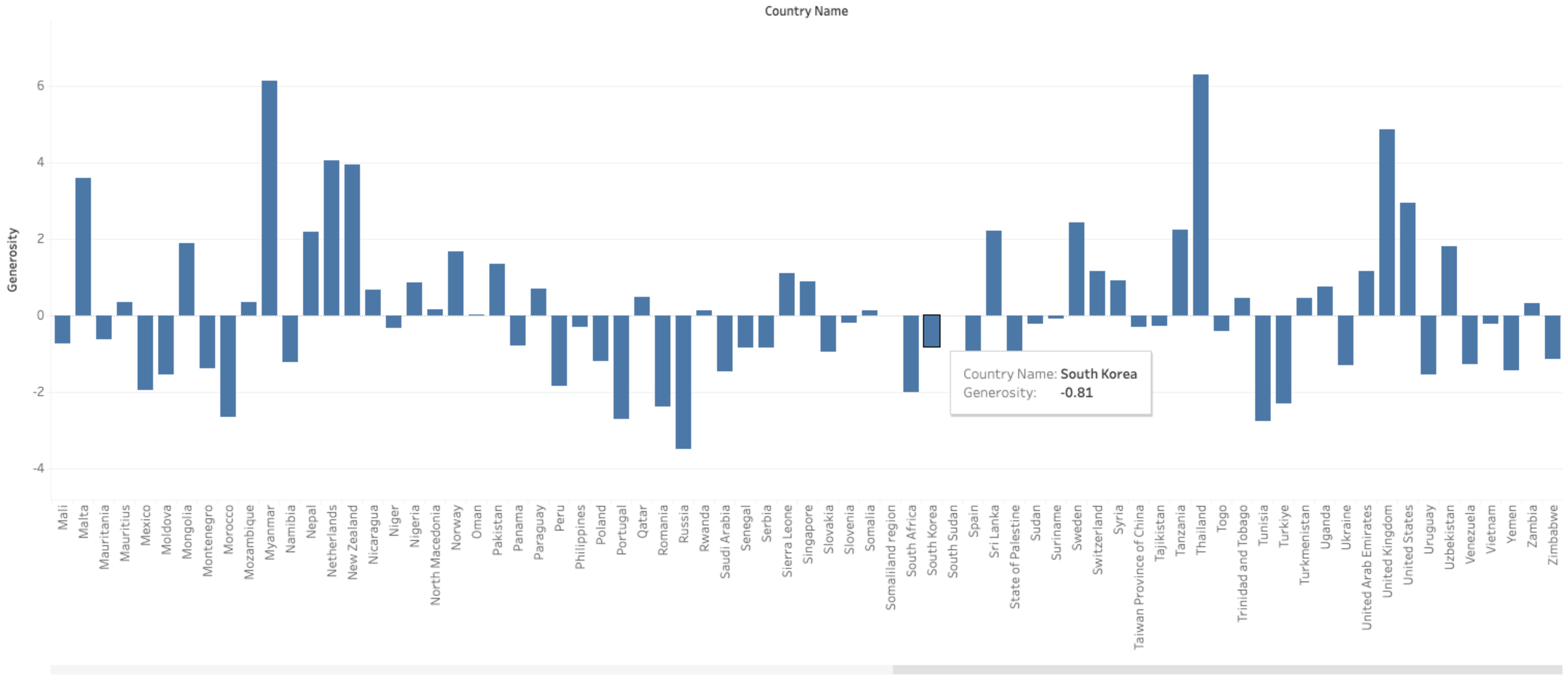
dt32 - Area Map Freedom



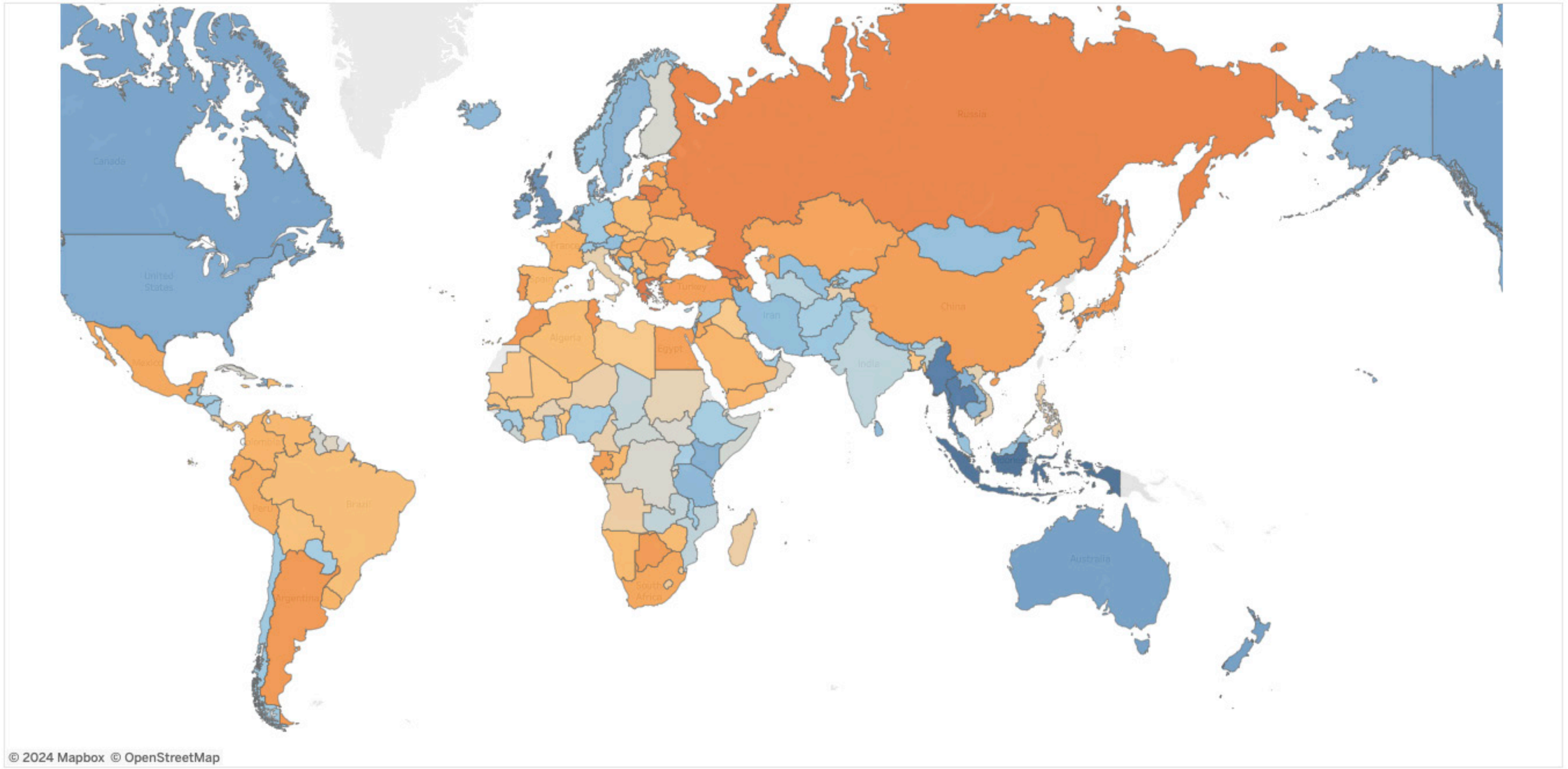
dt35 - Pie Chart



dt37 - Bar Chart 03 Generosity



dt37 - Area Map Generosity



Thank You