

State of Children's Education 2021

Education Accessibility around the World

Objective



Analyze the world's state of children's education using out of school and completion rate data. This analysis will provide insights into countries with significant education accessibility and gender gaps for further action.

Data



1. UNICEF's State of the World's Children 2021 on Education. Dataset can be downloaded from [Completion Rate](#), [Statistical tables](#), [Others](#).
2. GDP Per Capita data from World Bank. Dataset can be downloaded [here](#).
3. 2022 Global Happiness Rank by World Happiness Report. Dataset can be downloaded [here](#).

Tools & Skills

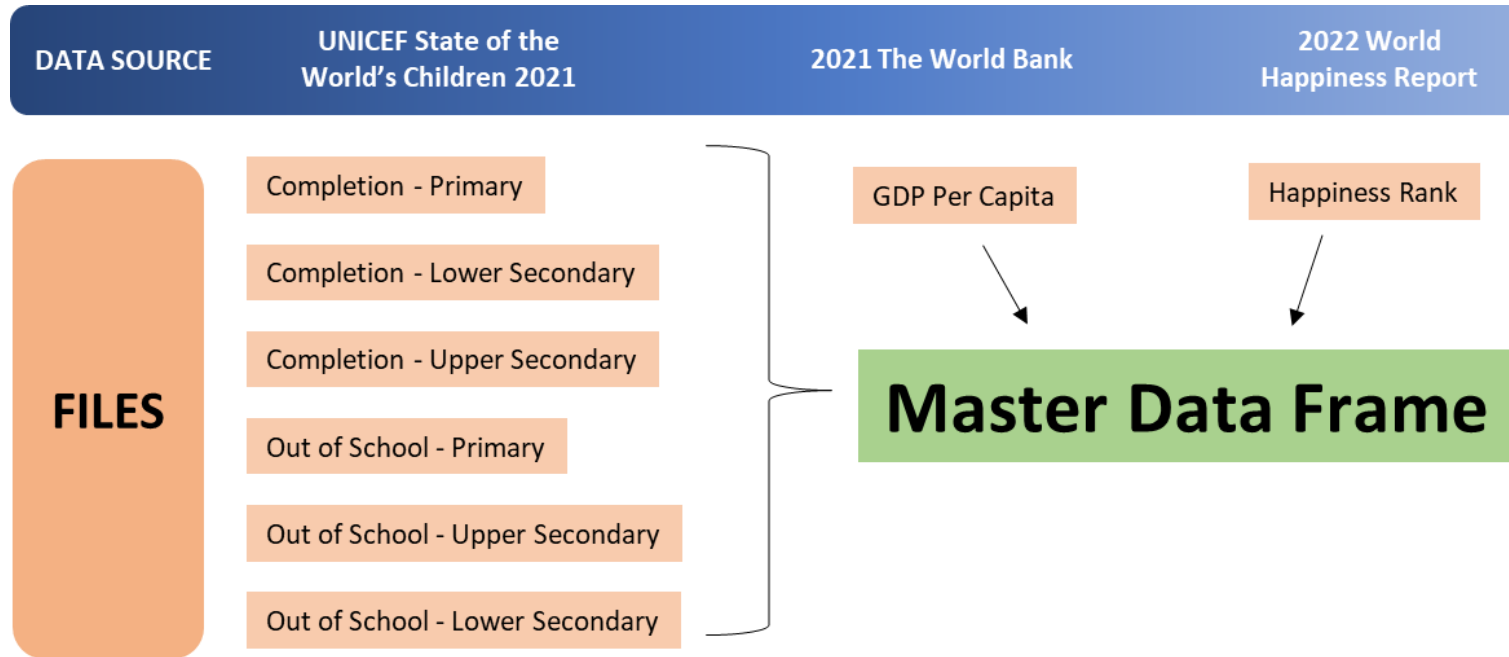


Data Exploring
Data Wrangling & Subsetting
Data Merging
Geographical Visualizations in Python
Supervised Machine Learning - Regression
Unsupervised Machine Learning - Clustering

Data Integration
Data Transformation

Visualizations & Forecasting
[Tableau Link](#)

Sourcing Data & Cleaning in Python



Cleaned and merged data from 8 separate files to produce a master dataframe

Challenges:

Standardizing country names every time data is merged from a new source and geojson file.

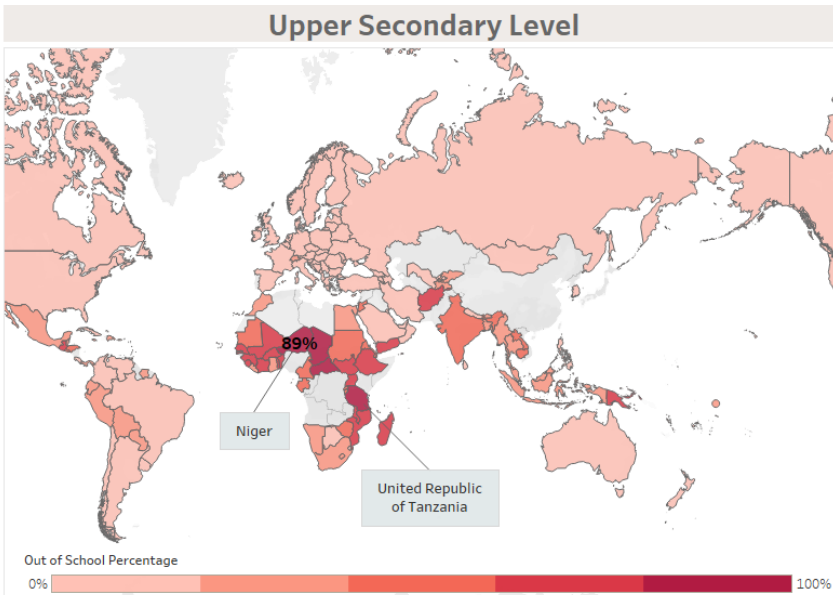
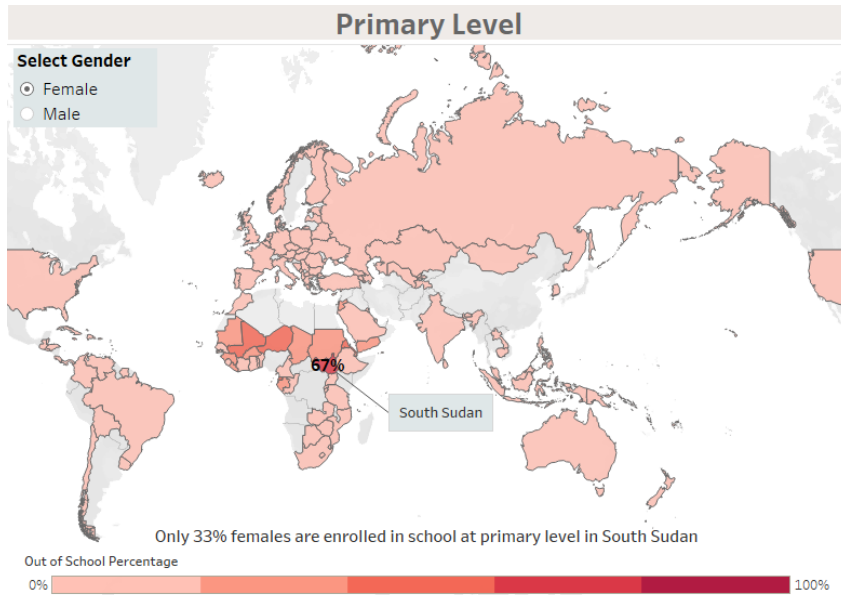
Solution: Create a reference country list based on UNICEF data. Conduct merge and look for left only or right only values and proceed to rename the countries not joined based on the reference list.

Cons: Does not work with geojson file. Had to rename to suit geojson file for visualization in Python. Working on solution that works by integrating geometry into main dataframe.

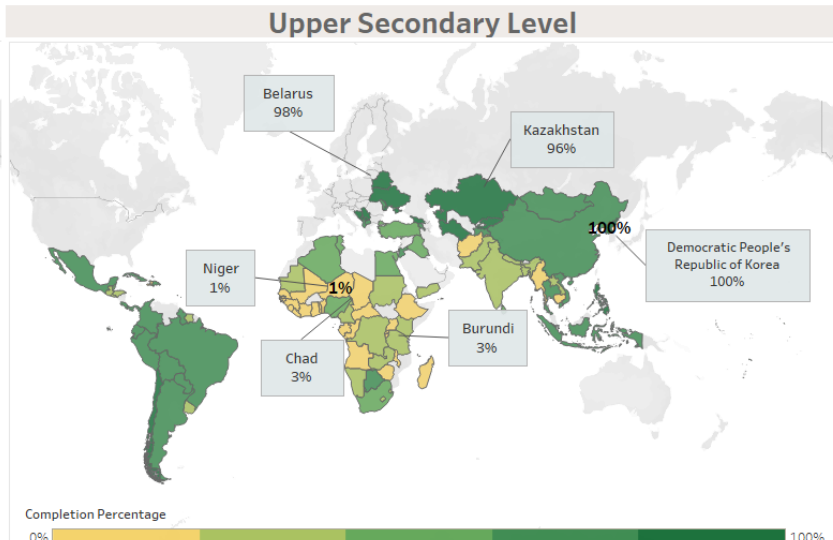
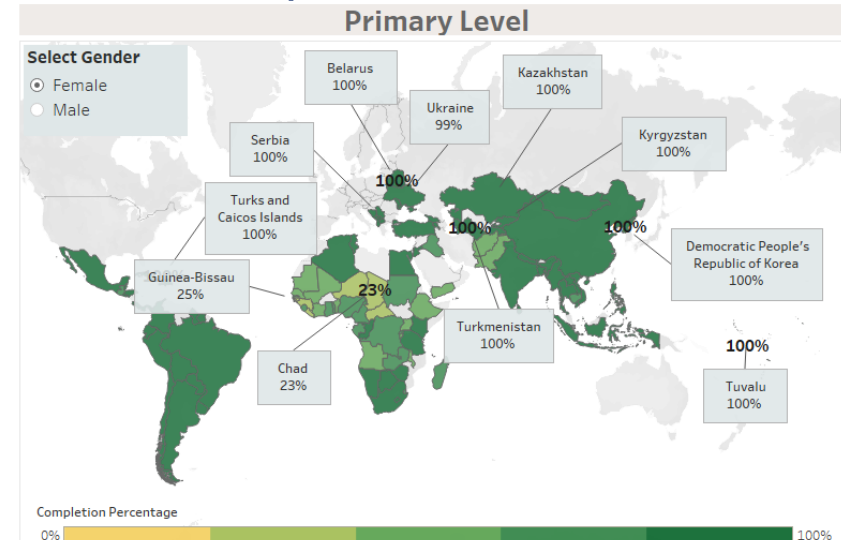
```
# renaming more countries in df to match with geojson.
df['Country'] = df['Country'].replace({
    "Bolivia (Plurinational State of)": "Bolivia",
    "Brunei Darussalam": "Brunei",
    "Côte d'Ivoire": "Ivory Coast",
    "Congo": "Republic of Congo",
    "Cabo Verde": "Cape Verde",
    "Czechia": "Czech Republic",
    "Micronesia (Federated States of)": "Federated States of Micronesia",
    "Guinea-Bissau": "Guinea Bissau",
    "Iran (Islamic Republic of)": "Iran",
    "Republic of Korea": "South Korea",
    "Lao People's Democratic Republic": "Laos",
    "Republic of Moldova": "Moldova",
    "Democratic People's Republic of Korea": "North Korea",
    "State of Palestine": "Palestine",
    "Russian Federation": "Russia",
    "Serbia": "Republic of Serbia",
    "Timor-Leste": "East Timor",
    "Türkiye": "Turkey",
    "United States": "United States of America",
    "Holy See": "Vatican",
    "Venezuela (Bolivarian Republic of)": "Venezuela",
    "Viet Nam": "Vietnam",
})
```

Findings

Out of school Rate:



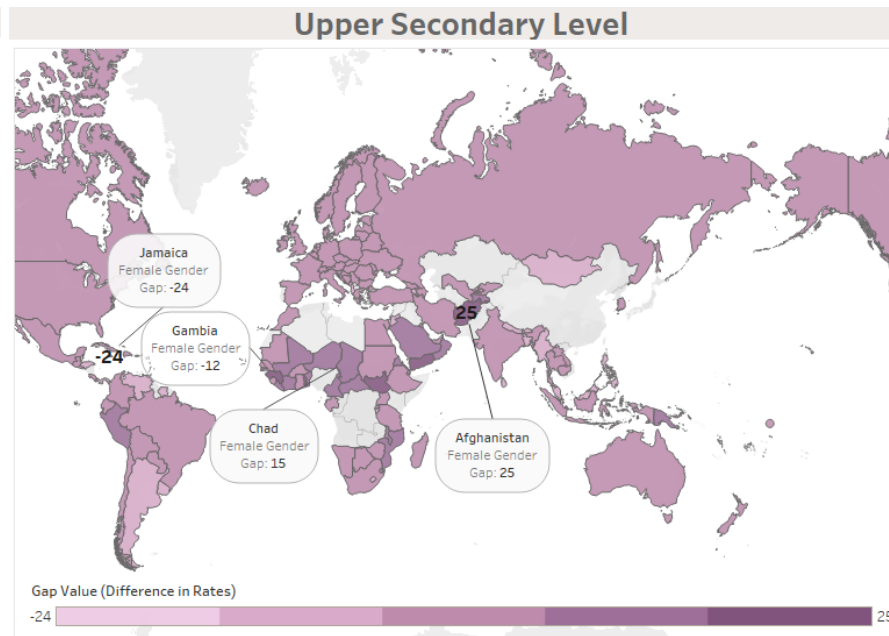
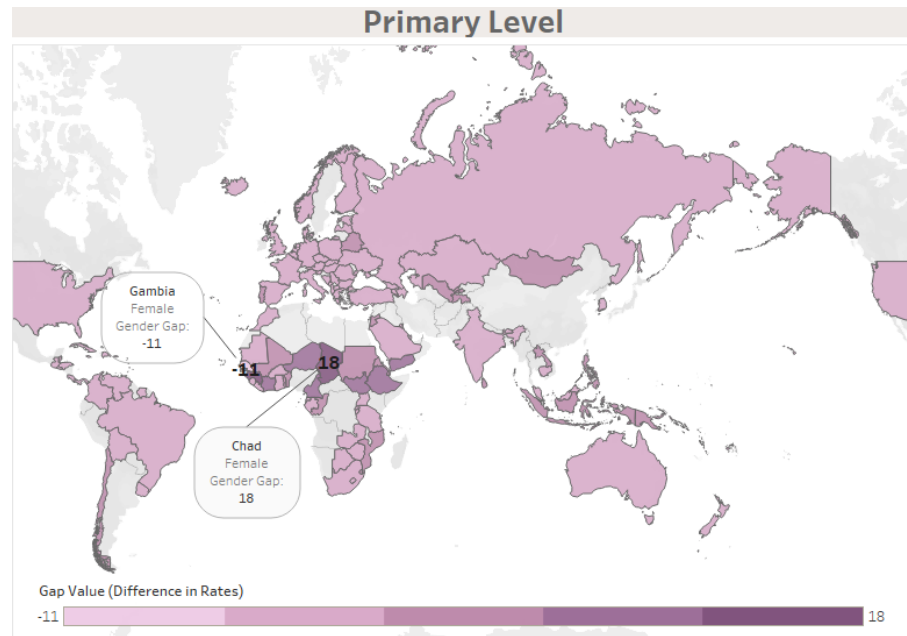
Education Completion Rate:



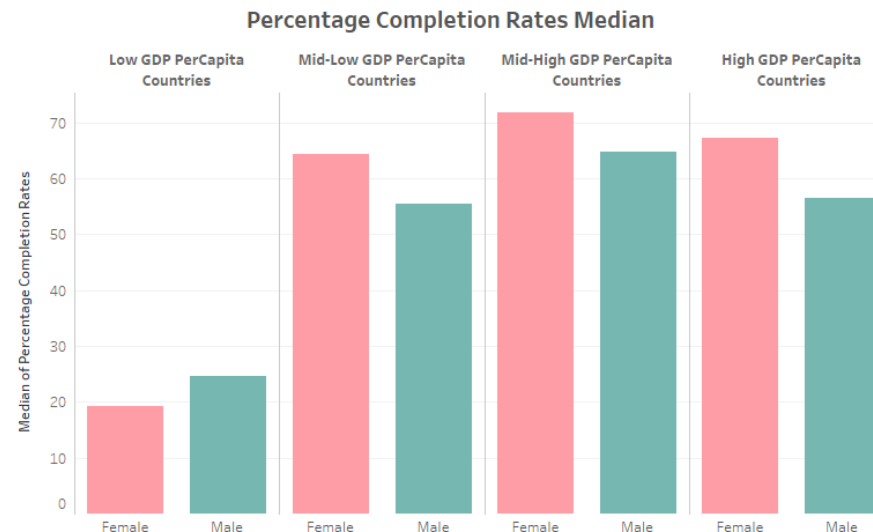
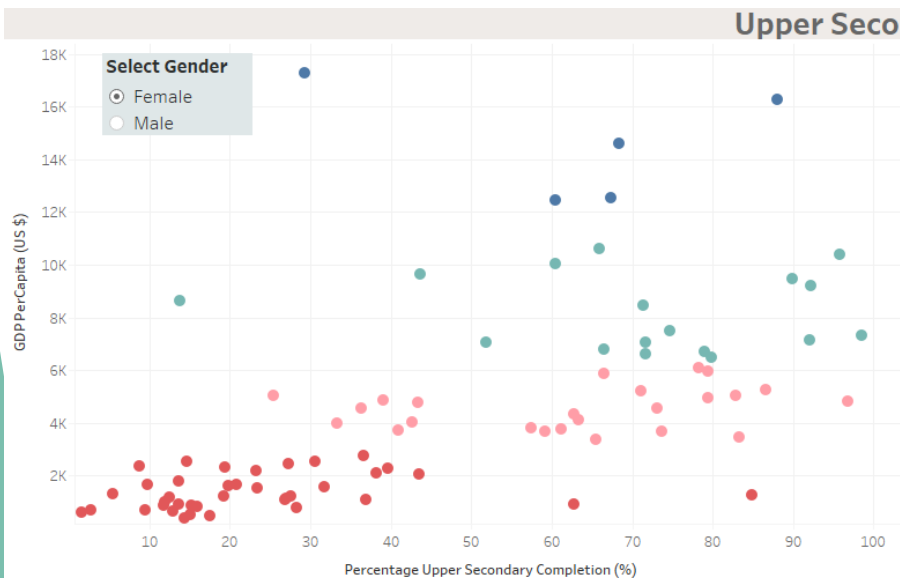
- In South Sudan, 67% of female primary school aged children are not in school.
- Those who managed to be in school, still find it hard to complete their studies, for example Chad having as low as only 23% females completing it at primary level.
- As high as 89% of female upper secondary students in Niger did not complete their studies.

(Dynamic visualization on [Tableau](#))

Gender Gap Analysis and Cluster Analysis



- The out of school rates can give insights into how big of a gap between male and female who are in school. Afghanistan has the biggest female gap in the upper secondary level and Chad has the biggest gap for primary level.
- Education completion rates do impact the GDP as a measure of the country's development, evidently in the cluster analysis.
- Pushing for gender equality and access to education for females might benefit the country.



(Dynamic visualization on [Tableau](#))

Next Steps



Run cluster analysis on Happiness Rank vs completion rates and examine relationships.



Deep dive into specific countries for further analysis on time series and forecasting if required

