Pollution and Prosperity

Studying the link between economics and air quality

By Gabriella Weis and Mason Polier

Background

- **Two datasets:**
  - **GDP** - World Development Indicators (World Bank)
  - **Air Quality** - Global Air Pollution (Kaggle)

- **Variables of interest:**
  - AQI Category, AQI Value
  - Country GDP (billions USD)
  - Agriculture, Industry, Manufacturing, Services (% of GDP allocated)

Gross domestic product		Agriculture		Industry		Manufacturing		Services	
$ billions		% of GDP		% of GDP		% of GDP		% of GDP	
2015	2023	2015	2023	2015	2023	2015	2023	2015	2023
19.1	14.5	20.6	33.7	22.1	16.1	11.4	10.2	53.2	45.0
11.4	23.0	19.8	18.3	21.8	21.2	5.7	5.9	46.3	48.0
187.5	239.9	10.5	13.2	32.8	38.0	7.1	7.8	52.1	45.1
0.7	0.9	..	..	..	..	14.0	8.7	..	..
2.8	3.7	0.5	0.5	10.0	11.4	3.7	3.6	78.7	78.6
90.5	84.7	9.1	14.9	41.9	45.3	5.7	8.0	48.7	39.7
1.4	2.0	1.4	1.9	15.9	19.4	2.8	2.5	73.1	67.8
594.7	640.6	5.2	6.1	23.2	25.1	14.2	16.3	55.8	52.8

Snapshot from World Bank Dataset

Process

	Country	GDP	Agriculture	Industry	Manufacturing	Services	City
1	Afghanistan	14.5	33.7	16.1	10.2	45.0	Kuhestan
2	Afghanistan	14.5	33.7	16.1	10.2	45.0	Qunduz
3	Afghanistan	14.5	33.7	16.1	10.2	45.0	Rostaq
4	Afghanistan	14.5	33.7	16.1	10.2	45.0	Tokzar
5	Afghanistan	14.5	33.7	16.1	10.2	45.0	Carikar
6	Afghanistan	14.5	33.7	16.1	10.2	45.0	Uruzgan
7	Afghanistan	14.5	33.7	16.1	10.2	45.0	Herat
8	Afghanistan	14.5	33.7	16.1	10.2	45.0	Baglan
9	Afghanistan	14.5	33.7	16.1	10.2	45.0	Andarab
10	Afghanistan	14.5	33.7	16.1	10.2	45.0	Nahrin

*Snapshot of joined data in R*

- **Difficulties with the data**
  - GDP dataset = based around countries
  - Air Quality dataset = based around cities
  - Result = many observations per country and other issues
  - USA had no Air Quality data
- **Questions and exploration**
  - First - identified what we wanted to know and selected data accordingly
  - Second - created questions
 - Third - eliminated any questions that were unfeasible

Research Questions:

- What is the distribution of cities across AQI categories in countries with a GDP over $1 trillion? How does this compare with countries with a GDP under $1 trillion?
- How does the average AQI category vary among countries based on their primary GDP sector allocation?
- How does GDP percentage vary across sectors?
- Is there a correlation between overall GDP and a country's average AQI value?
- How do the countries with the top 5 GDPs compare in terms of air quality?
- How do average AQI values differ between countries that allocate the maximum percentage of their GDP to industry versus those that allocate the maximum to another sector?
- How does PM2.5 AQI category differ between cities in countries that allocate a higher percentage of their GDP to agriculture (over 20%) versus those that allocate a lower percentage?

Q#1 What is the distribution of cities across AQI categories in countries with a GDP over $1 trillion? How does this compare with countries with a GDP under $1 trillion?
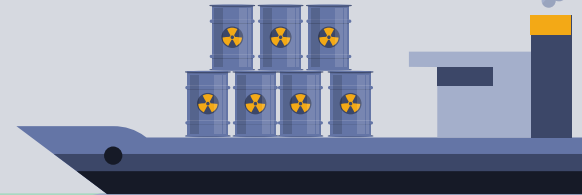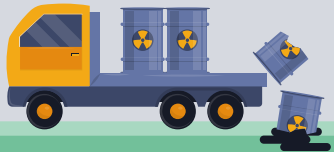
Cities in Countries with a GDP Over $1 Trillion Across AQI Categories

Cities in Countries with a GDP Under $1 Trillion Across AQI Cats.
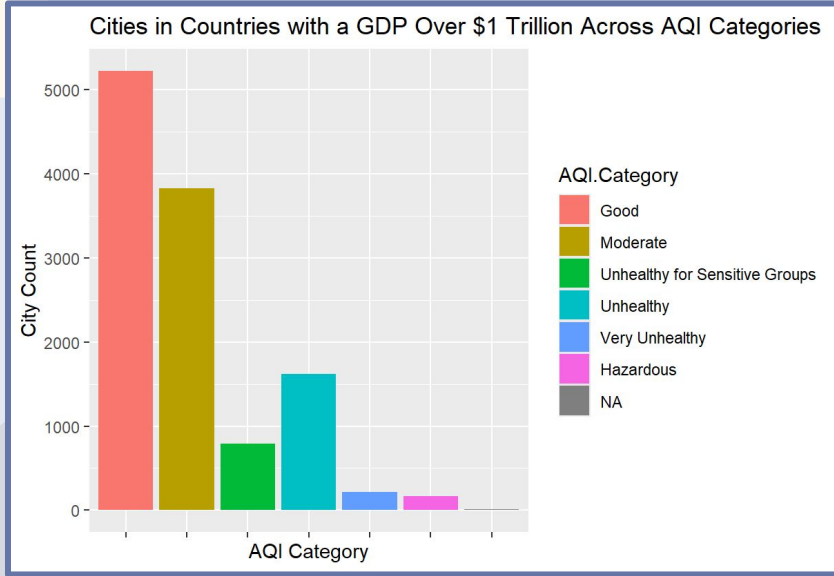
Majority of cities = Good
Higher counts of Unhealthy than for USG
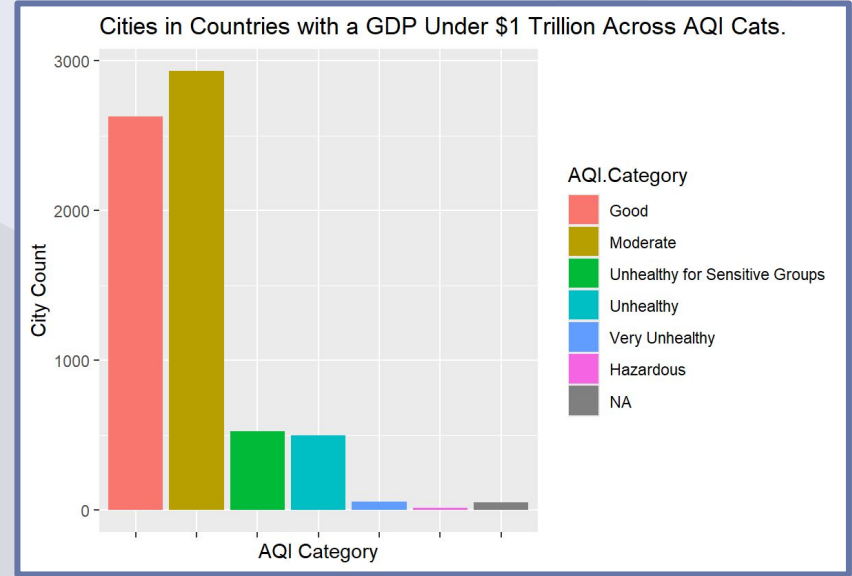
Majority of cities = Moderate
Higher counts of USG than for Unhealthy

The percentages of cities with Hazardous and Unhealthy ratings appear to be lower in countries with a GDP under $1 trillion than those in countries with a GDP over $1 trillion.

Q#2 How does the average AQI category vary among countries based on their primary GDP sector allocation?

Average AQI vs. Max. Sector for Countries with a GDP Over $1 Trillion
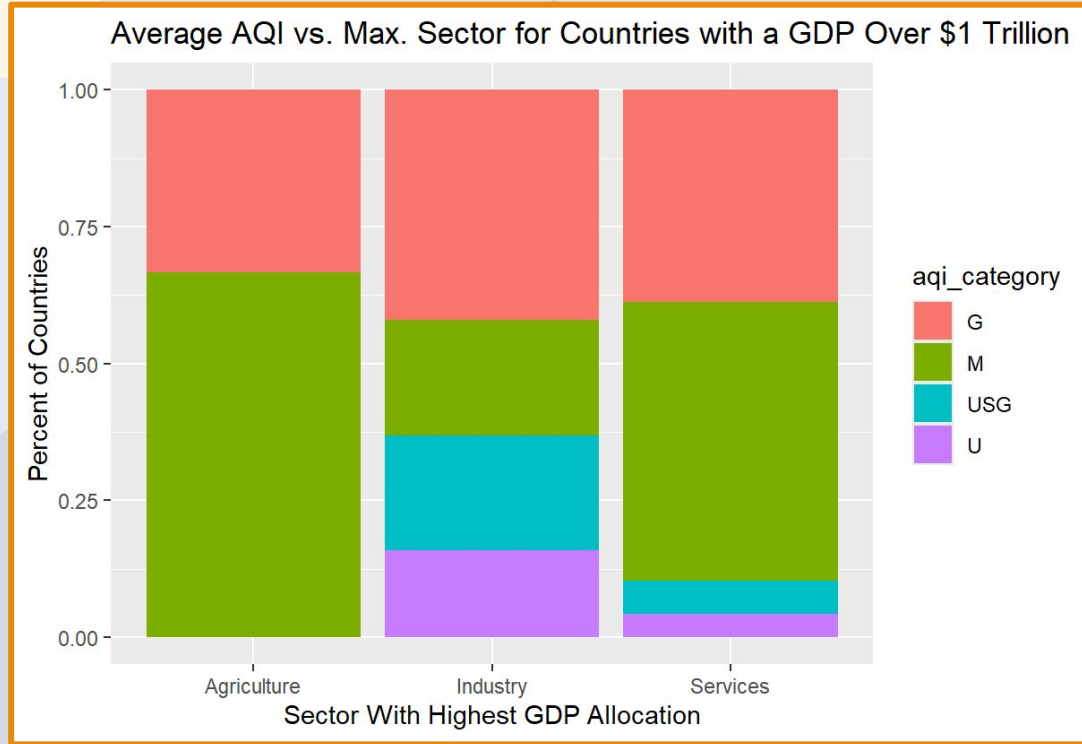
Creation:
- New variables
  - Max Sector (categorical)
  - Average AQI Category (categorical)
- Remove NAs

Analysis:
- Over 50% of countries with a Max Sector of agriculture have an average AQI rating of Moderate
- Compared to other countries, those that allocate most of their GDP to Industry have more Unhealthy ratings

How does GDP percentage vary across sectors?

Creation:
- Grouped by sector
- GDP Percentage = percentage of country GDP allocated to the sector

Analysis:
- Largest IQR and median = Services
- Smallest IQR = Manufacturing
- Smallest median = Agriculture
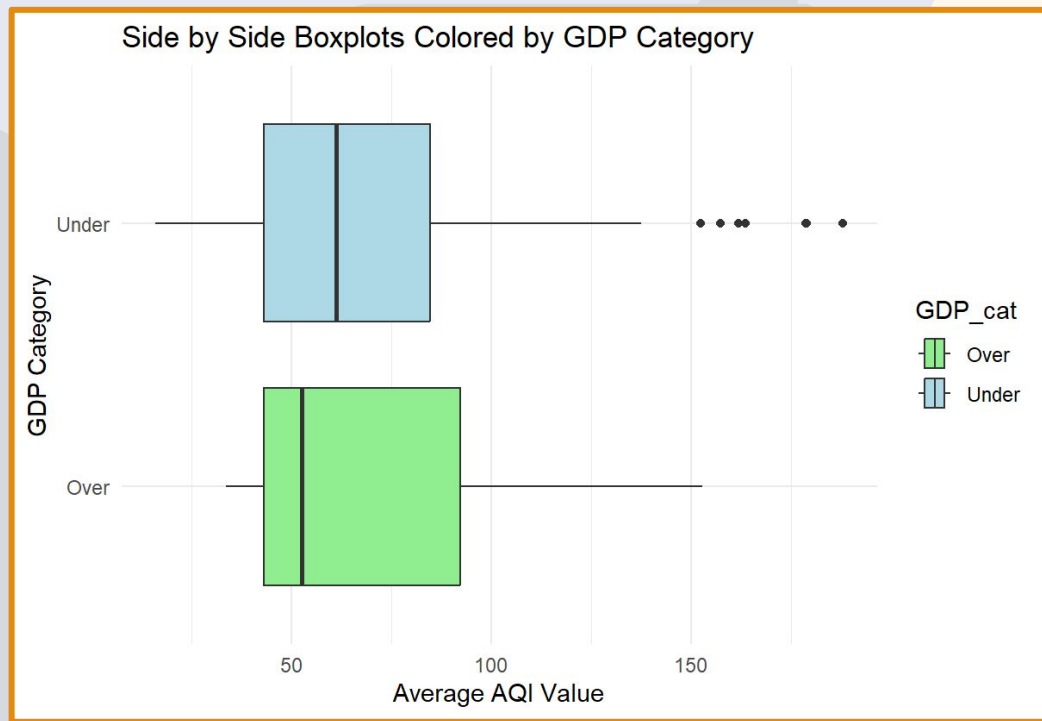- Services = greatest % of GDP allocation among countries

Distribution of GDP Percentages by Sector
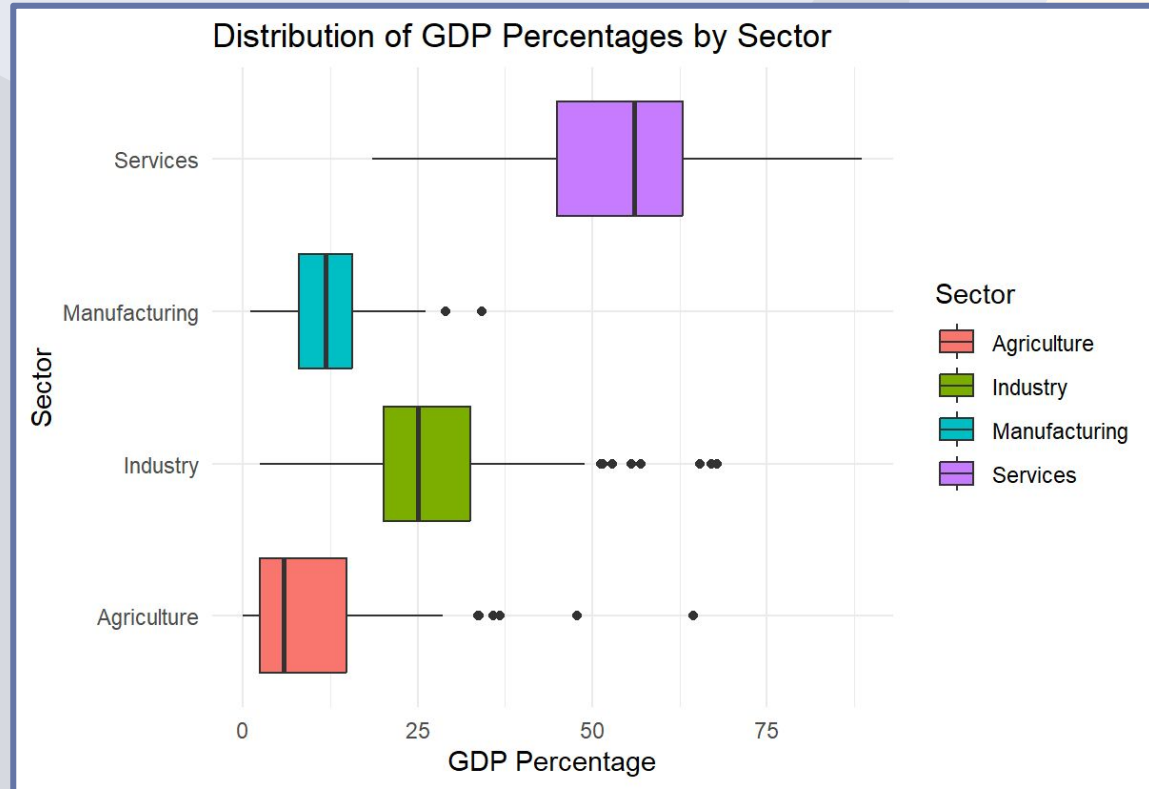
Q#4   Is there a correlation between overall GDP and a country's average AQI value?

Creation:
- New variable
  - GDP_cat (categorical)
  - Average AQI Value (numeric)
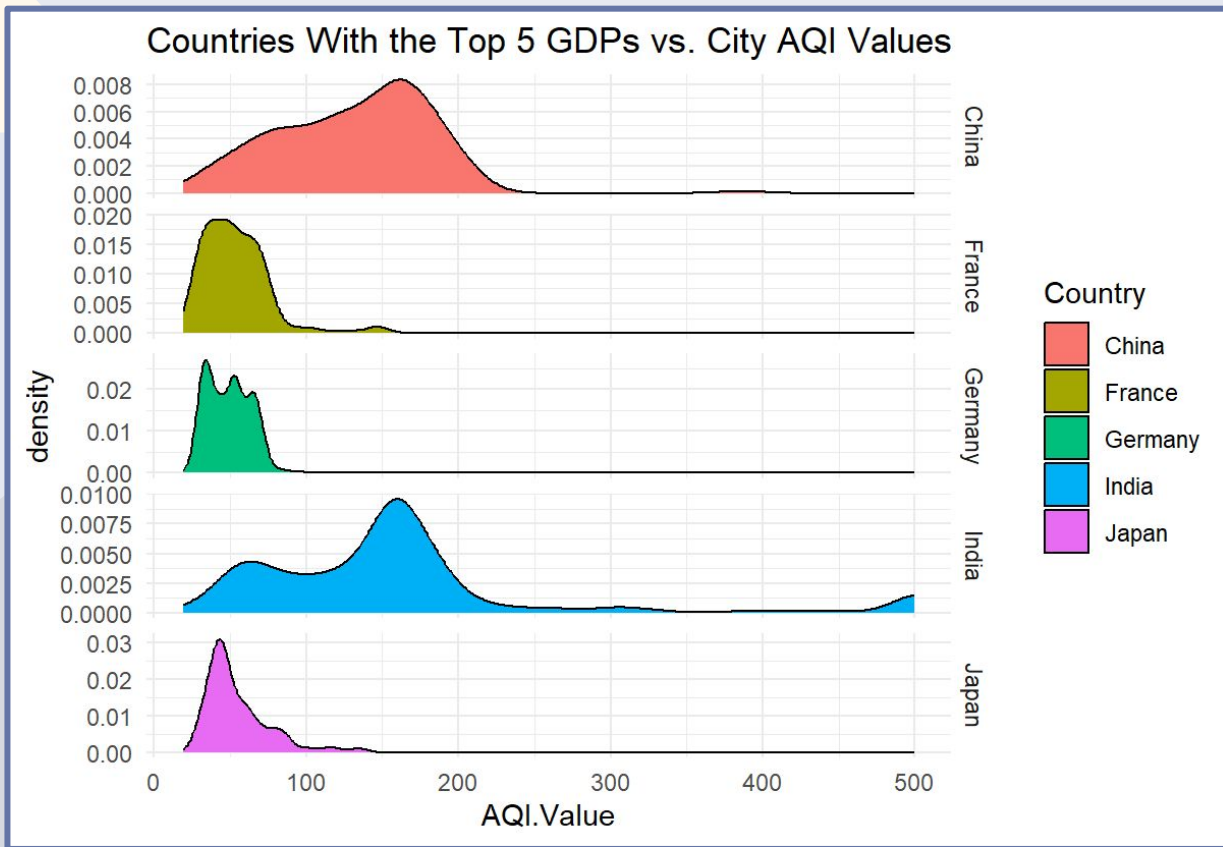- Remove NAs

Analysis:
- Over:
  - Median = 52.53
  - Larger IQR range (49.14)
  - Outliers? - no
- Under:
  - Median = 61.20
  - Smaller IQR range (41.57)
  - Outliers? - yes, multiple

Side by Side Boxplots Colored by GDP Category

How do the countries with the top 5 GDPs compare in terms of air quality?

Countries With the Top 5 GDPs vs. City AQI Values
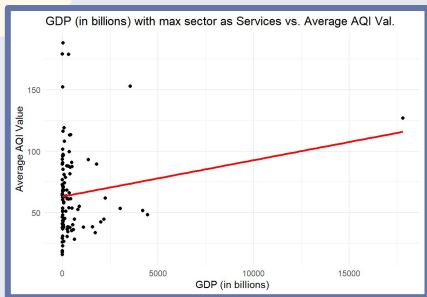
Creation:
- Filter out countries with the top 5 GDPs (notice the absence of USA)
- Facet grid plots = distributions of cities' AQI values across countries
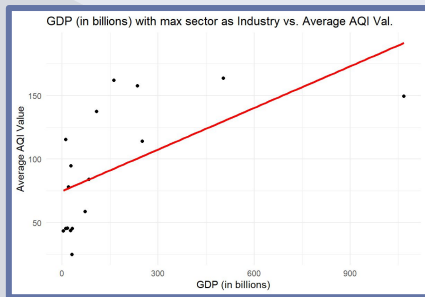
Analysis:
- India = largest median and IQR
- Japan = smallest median and IQR
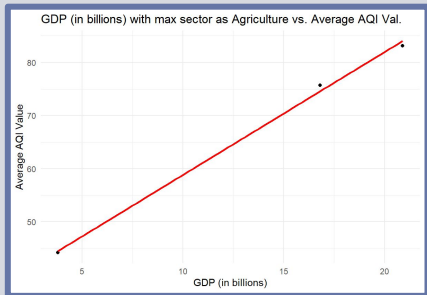- Allows us to see the variety between countries

How do average AQI values differ between countries that allocate the maximum percentage of their GDP to industry versus those that allocate the maximum to another sector?
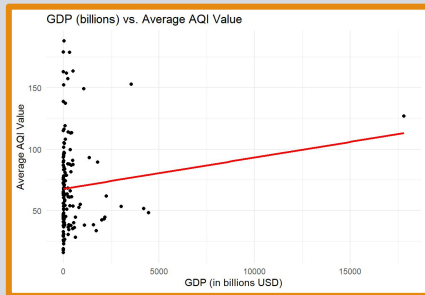

GDP (in billions) with max sector as Services vs. Average AQI Val.

R = 0.1617831

GDP (in billions) with max sector as Industry vs. Average AQI Val.

R = 0.601604

GDP (in billions) with max sector as Agriculture vs. Average AQI Val.

R = 0.9987477

GDP (billions) vs. Average AQI Value
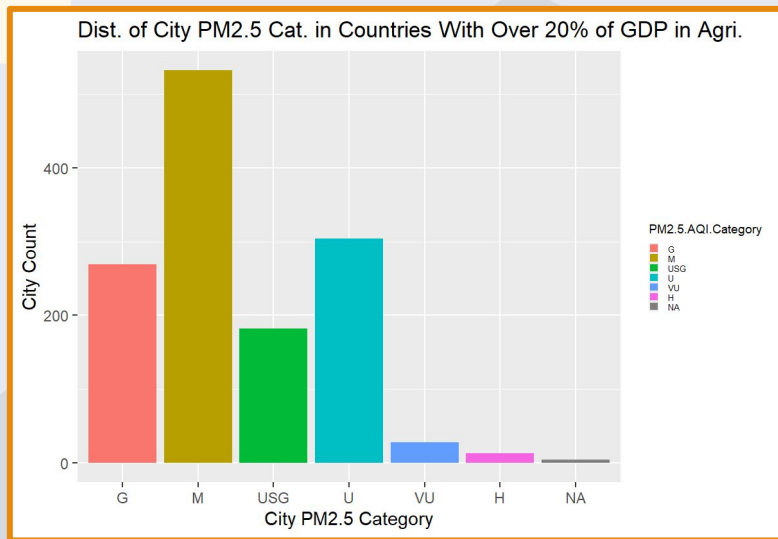
R = 0.1123833

Creation:
- Filtered by Max Sector
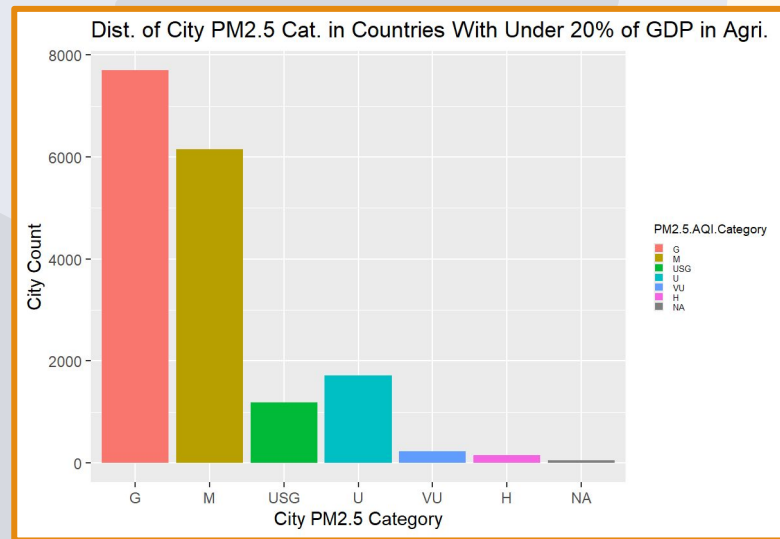- Set Average AQI Value against each sector

Analysis:
- Industry:
  - moderately strong, positive, linear correlation
- Services:
  - very weak, positive, linear correlation
- Agriculture:
  - very strong, positive, linear correlation

How does PM2.5 AQI category differ between cities in countries that allocate a higher percentage of their GDP to agriculture (over 20%) versus those that allocate a lower percentage?

Dist. of City PM2.5 Cat. in Countries With Over 20% of GDP in Agri.

Dist. of City PM2.5 Cat. in Countries With Under 20% of GDP in Agri.

Good = 20.2% of cities
Moderate = 39.94% of cities
Lower number of cities than the other graph

Good = 44.85% of cities
Moderate = 35.82% of cities
Higher number of cities than the other graph

While cities with over 20% of their country's GDP allocated to Agriculture tend to have a higher chance of being in the Moderate AQI category and a lower chance of being in the Good AQI category, cities with under 20% of their country's GDP allocated to Agriculture tend to have the opposite chances.

Conclusions & Future Work

- **Major conclusions:**
    - Hazardous and Unhealthy ratings vs. GDP
    - Services = greatest % of allocation
    - Industry vs. Average Country AQI Value
    - Possible Simpson's Paradox
    - Agricultural allocation vs. AQI categorization
- **Future:**
    - Working with increasing GDP vs. AQI
    - Data that includes countries that might have a large impact on our conclusions

Thank You!

Questions?