Fuel consumption data: Look for data on fuel consumption by different types of vehicles, such as cars, motorcycles, buses, and trucks. This data will help you understand the amount of fuel consumed by different vehicle categories and their contribution to emissions.

Vehicle registration data: Analyze data on the number of registered vehicles in different regions or cities. This will give you insights into the vehicle density and the potential impact on emissions.

Vehicle emission standards data: Explore data on vehicle emission standards, such as Bharat Stage (BS) standards in India. This information will help you assess the environmental performance of vehicles and identify areas where improvements can be made.

Traffic volume data: Gather data on traffic volume, including the number of vehicles passing through specific road segments or intersections. This data can be used to identify congested areas that may contribute to increased emissions.

Air quality data: Obtain data on air quality measurements, including levels of pollutants such as carbon monoxide (CO), nitrogen oxides (NOx), particulate matter (PM), and volatile organic compounds (VOCs). By comparing this data with other relevant factors, you can determine the impact of road transport emissions on air quality.

Geospatial data: Utilize geospatial data, such as maps and satellite imagery, to visualize the distribution of vehicles, traffic patterns, and areas with high emissions. This can help identify hotspots or regions that require immediate attention.

Once you have gathered the necessary data, you can perform various analyses to identify the biggest contributors to emissions and suggest mitigation strategies. These analyses can include:

Aggregating fuel consumption data by vehicle category to determine which types of vehicles contribute the most to emissions.

Correlating vehicle registration data with emission standards data to identify regions with a higher density of older or less efficient vehicles.

**Choosing data**

Analyzing traffic volume data to identify congested areas and recommend traffic management measures to reduce emissions.

Comparing air quality data with other relevant factors, such as traffic volume and vehicle density, to identify areas with significant emission-related air pollution.

Visualizing the geospatial data to identify areas where interventions, such as improved public transportation or cycling infrastructure, could help reduce emissions.

For the code walkthrough, you can use programming languages like Python or R to clean, analyze, and visualize the data. You can use libraries such as Pandas, NumPy, Matplotlib, or ggplot for data manipulation and visualization.

During the presentation, you can showcase the insights gained from the data analysis, including the biggest contributors to emissions and potential measures to tackle them. Use visuals such as graphs, charts, and maps to illustrate your findings effectively. Additionally, consider presenting the limitations of the analysis and potential areas for further research or data collection.