

Dataset Description

1 DESCRIPTION:

- **Size of the dataset:** 9358 instances
- **Number of attributes:** 15 attributes
- **Name and type of attributes:**

i) Date (Date)	---	>	Nominal
ii) Time (Time)	---	>	Nominal
iii) CO(GT) (Gas Concentration in mg/m ³)	---	>	Continuous
iv) PT08.S1(CO) (Sensor CO concentration)	---	>	Continuous
v) NMHC(GT) (Non-Methane Hydrocarbons concentration)	---	>	Continuous
vi) C6H6(GT) (Benzene concentration in µg/m ³)	---	>	Continuous
vii) PT08.S2(NMHC) (Sensor NMHC concentration)	---	>	Continuous
viii) NOx (GT) (Nitrogen Oxides concentration)	---	>	Continuous
ix) PT08.S3(NOx) (Sensor NOx concentration)	---	>	Continuous
x) NO2(GT) (Nitrogen Dioxide concentration)	---	>	Continuous
xi) PT08.S4(NO2) (Sensor NO2 concentration)	---	>	Continuous
xii) PT08.S5(O3) (Sensor Ozone concentration)	---	>	Continuous
xiii) T (Temperature in °C)	---	>	Continuous
xiv) RH (Relative Humidity in %)	---	>	Continuous
xv) AH (Absolute Humidity)	---	>	Continuous
- **Number of attributes according to types:**
 - (1) Qualitative (Nominal): 2
 - (2) Quantitative (Continuous): 13

2 POTENTIAL DATA MINING APPLICATION:

This dataset could be used for **predictive modeling** in air quality monitoring systems. For example, a regression model could predict pollutant concentrations based on sensor data to alert the public about air quality conditions. This would be useful in health-risk assessment applications, especially in urban areas where pollution levels can vary significantly throughout the day.