DAV MPA 1 Inferences Document

1. ) In this MPA Cancer Data Analysis and Visualisations are being done
2. The Data has below columns

Patient Id object

Age int64

Gender int64

Air Pollution int64

Alcohol use int64

Dust Allergy int64

OccuPational Hazards int64

Genetic Risk int64

chronic Lung Disease int64

Balanced Diet int64

Obesity int64

Smoking int64

Passive Smoker int64

Chest Pain int64

Coughing of Blood int64

Fatigue int64

Weight Loss int64

Shortness of Breath int64

Wheezing int64

Swallowing Difficulty int64

Clubbing of Finger Nails int64

Frequent Cold int64

Dry Cough int64

Snoring int64

Level object

1. The Analysis needs to be done with relation between various variables (features) of this data.
2. Usual Data Cleaning verifications are done.
3. Converted Categorical to and fro Numerical where necessary
4. Correlation between the variables is observed. The cancer level is impacted by various features listed at various levels

Air Pollution int64

Alcohol use int64

Dust Allergy int64

OccuPational Hazards int64

Genetic Risk int64

chronic Lung Disease int64

Balanced Diet int64

Obesity int64

Smoking int64

Passive Smoker int64

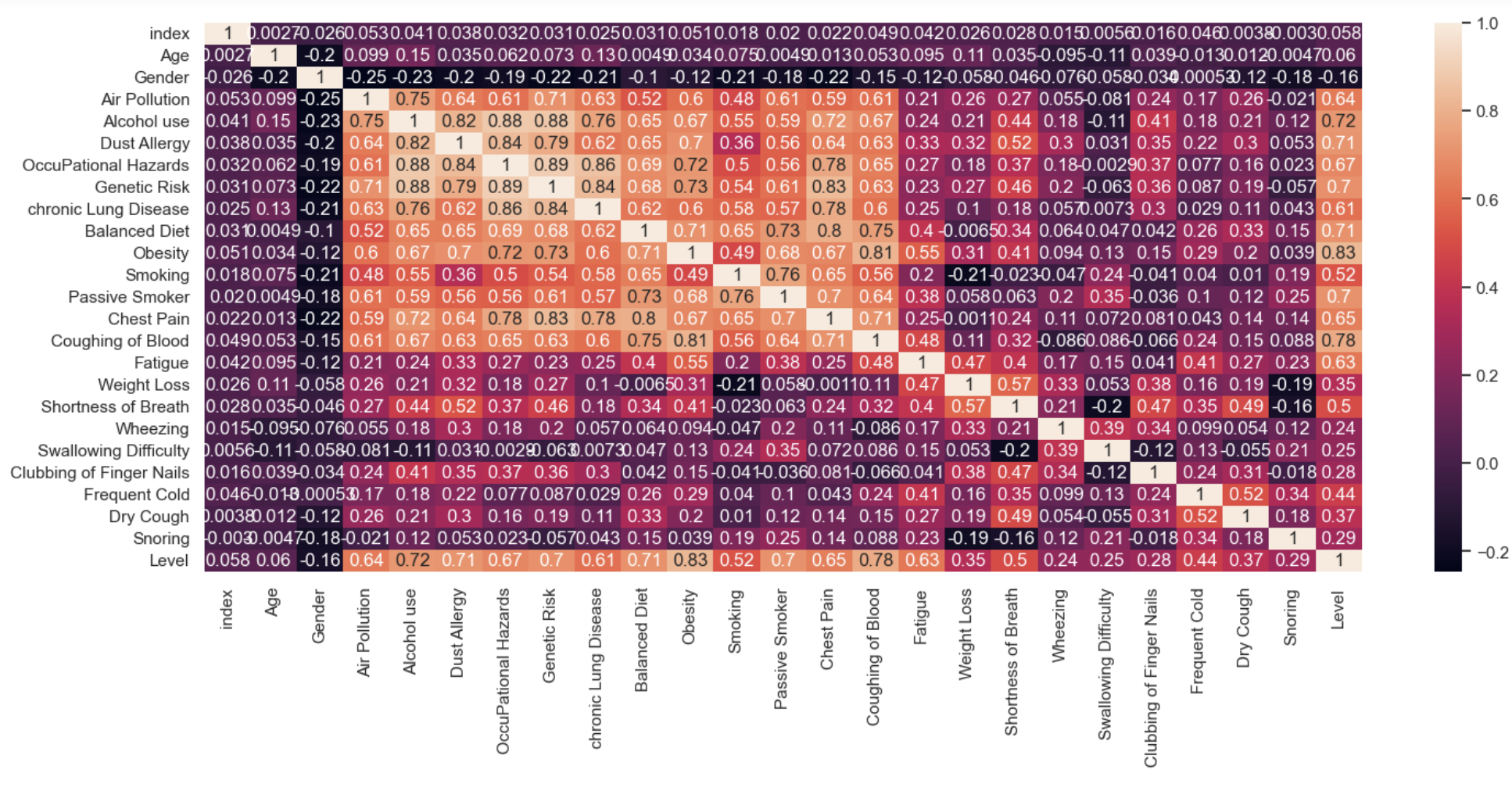
Chest Pain int64

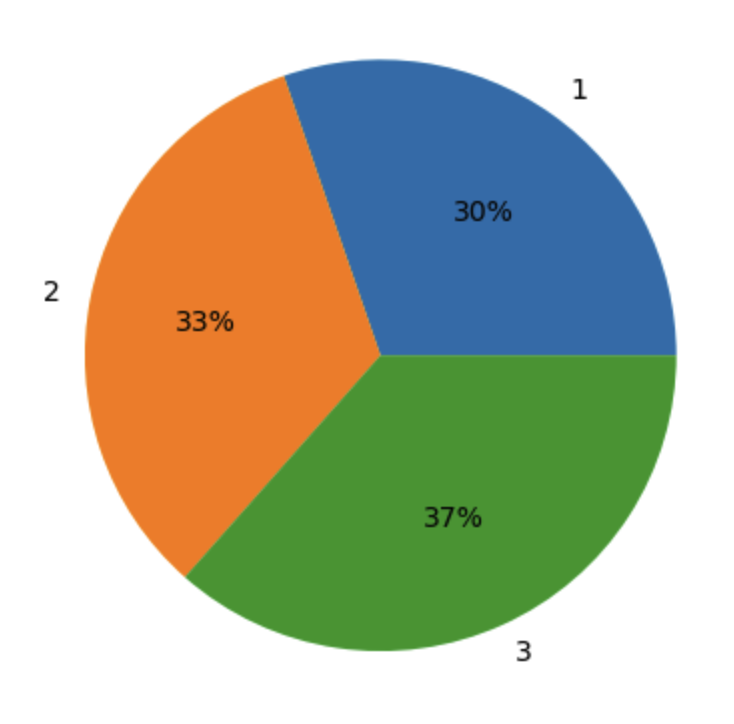
Coughing of Blood int64

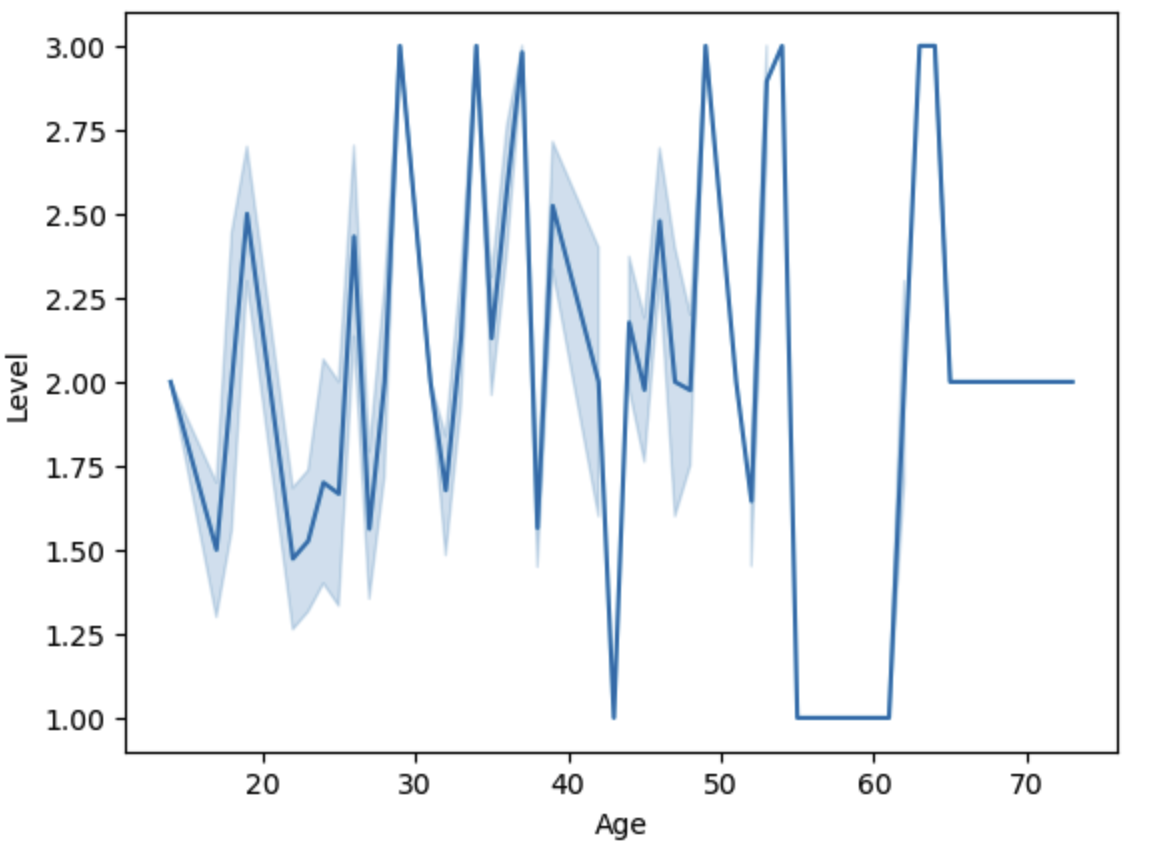
Fatigue int64

Weight Loss int64

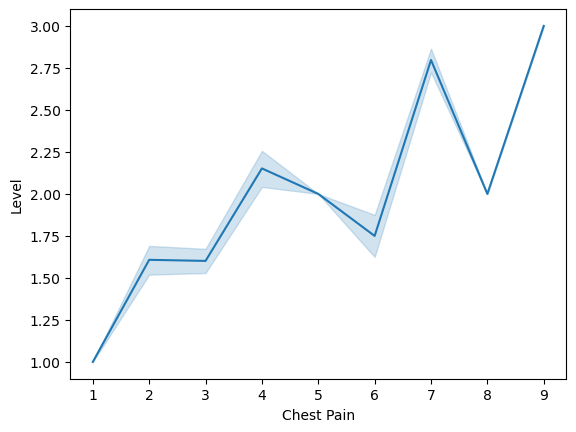
Shortness of Breath int64



1. Histogram depicts the similar observation
2. Number of Gender ‘1’ is observed to be high
3. Number of smokers who has cancer are more, especially in male
4. Cancer patients at different levels are almost equally distributed
5. Below line plot shows that Age is not a major parameter for cancer. Except that age at around 60 are having low level of cancer



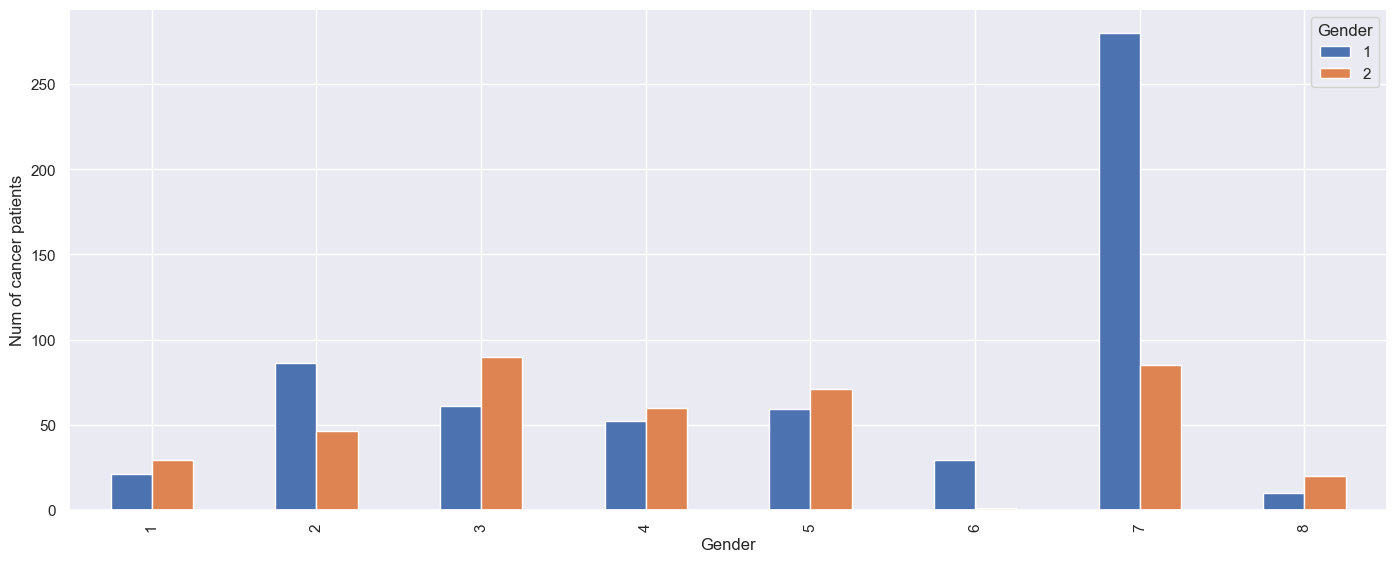
1. Below line plot evidences that people with more chest pain have more levels are cancer



13) Strong Positive correlation is observed between Smoking and shortness of breath

14) Age and Gender doesn’t show any correlation with cancer levels

15) Occupational hazards at level **7** has more number of cancer patients, for Gender1.



16) Obesity and Coughing of blood appears to be positively correlated as we can see that as the one increases second one also is increasing, excluding few exceptions.

