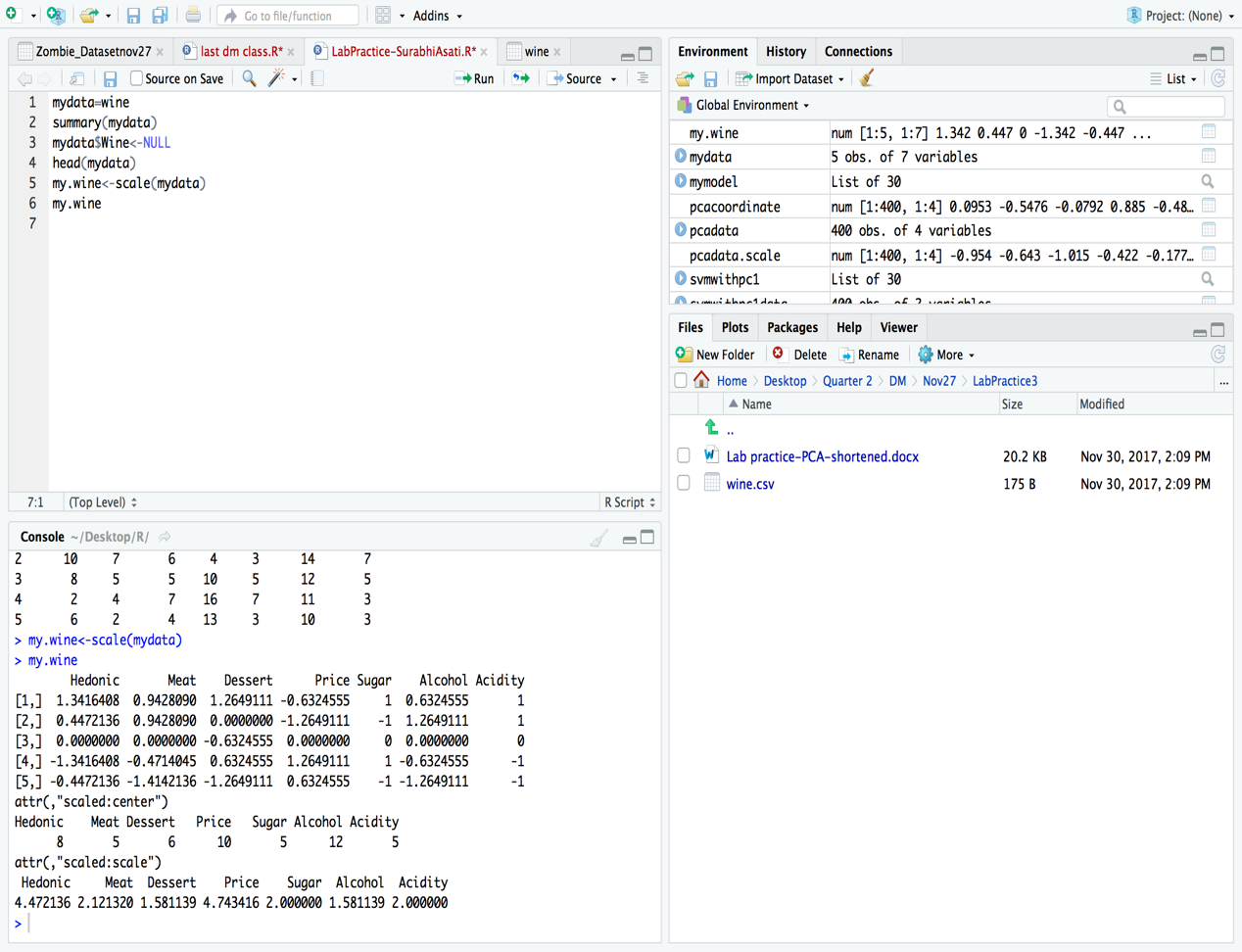
**Surabhi Asati**

**va7892**

ITM 6285 Data Mining Lab - Principal Component Analysis

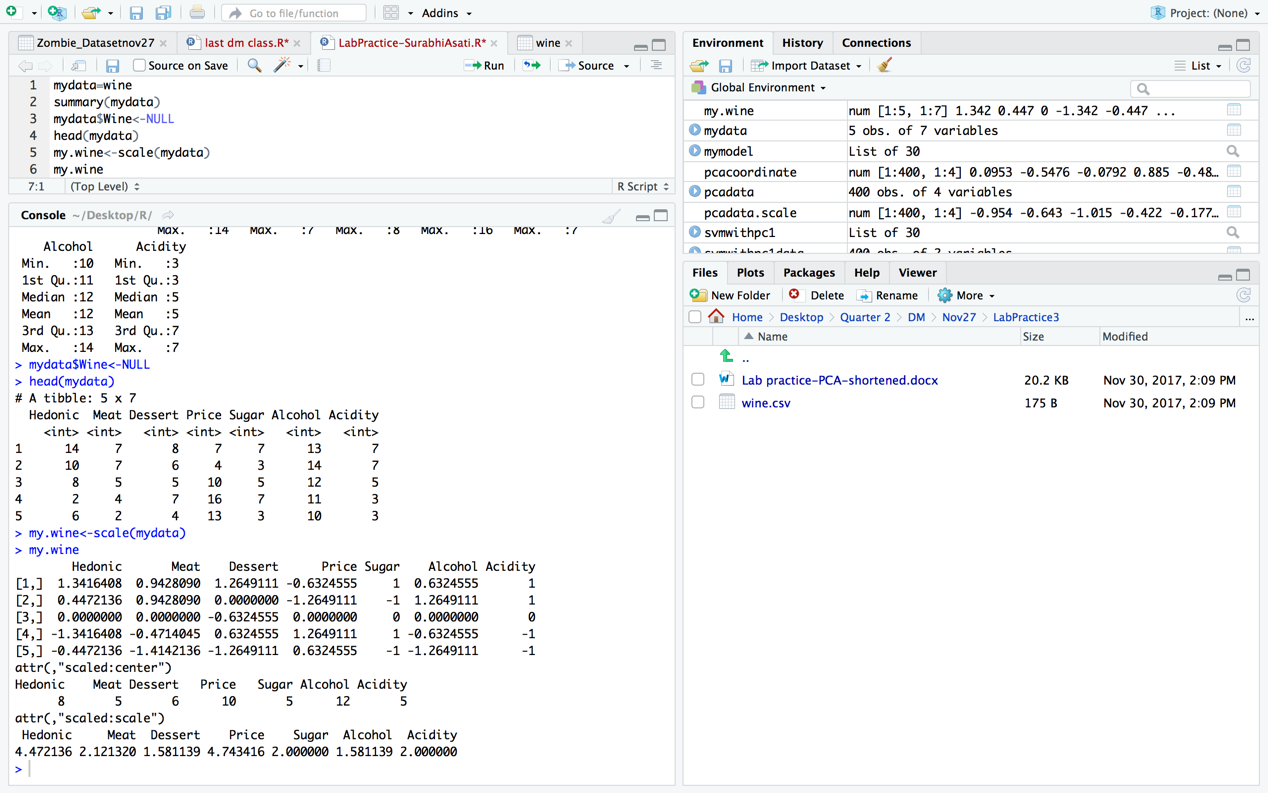
**Task 1: Import the data and Background Information**

One of the tasks is to find out the primary factors that distinguish between expensive wine and affordable wine.



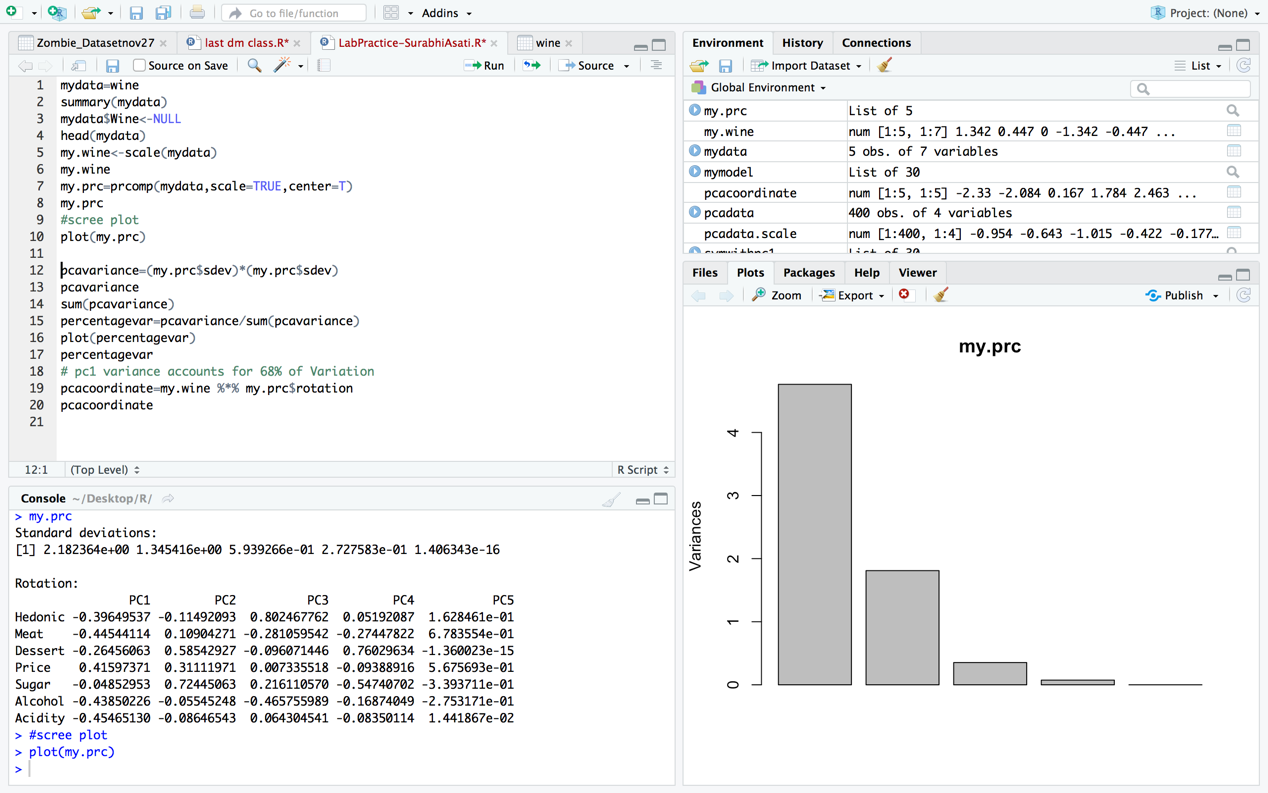
**Task 2: PCA using eigen value and eigen vector calculation-Data Preparation**

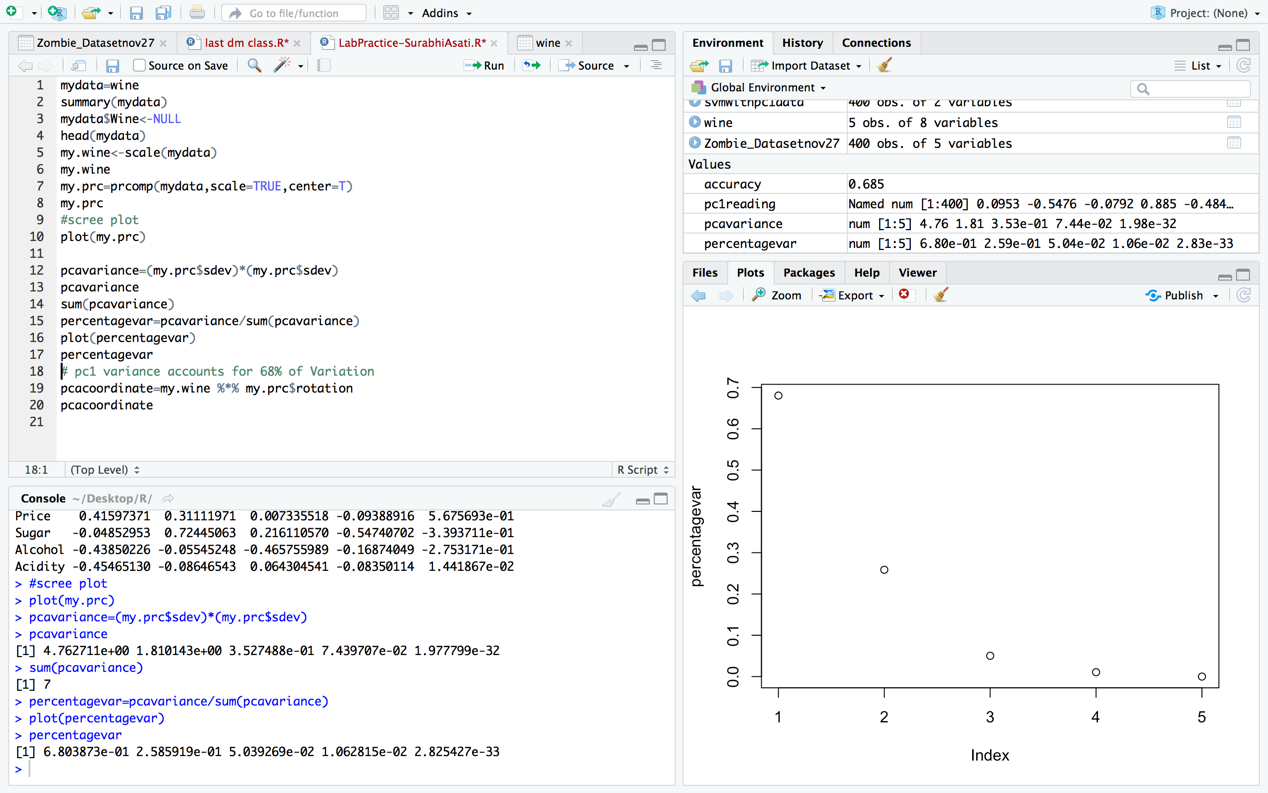
First, get rid of the wine name column, then *scale* the data, name the PCA-ready data set “my.wine”.



**Task 3: PCA using prcomp**

Do PCA using prcomp, show the R result and generate the screeplot, and then answer what is the primary contributor the PC2?





**Answer: The primary contributor is PC1 – 68% of variation**

**Price has the highest loading or Eigen value on PC1 with value = 0.41597371, Hence it is the primary contributor with 41.5%**

**Task 4: Calculate coordinates on principal components**

Calculate and copy the coordinates on the principal components to the answer sheet.

> pcacoordinate=my.wine %\*% my.prc$rotation

> pcacoordinate

PC1 PC2 PC3 PC4 PC5

[1,] -2.3301649 1.095284 0.67132194 0.09433756 -8.326673e-17

[2,] -2.0842419 -1.223185 -0.65633569 0.13366477 -8.326673e-17

[3,] 0.1673228 -0.370258 0.06076092 -0.48085363 8.601542e-16

[4,] 1.7842392 1.712563 -0.54923690 0.06463829 -3.608225e-16

[5,] 2.4628448 -1.214405 0.47348974 0.18821300 -2.498002e-16

>

