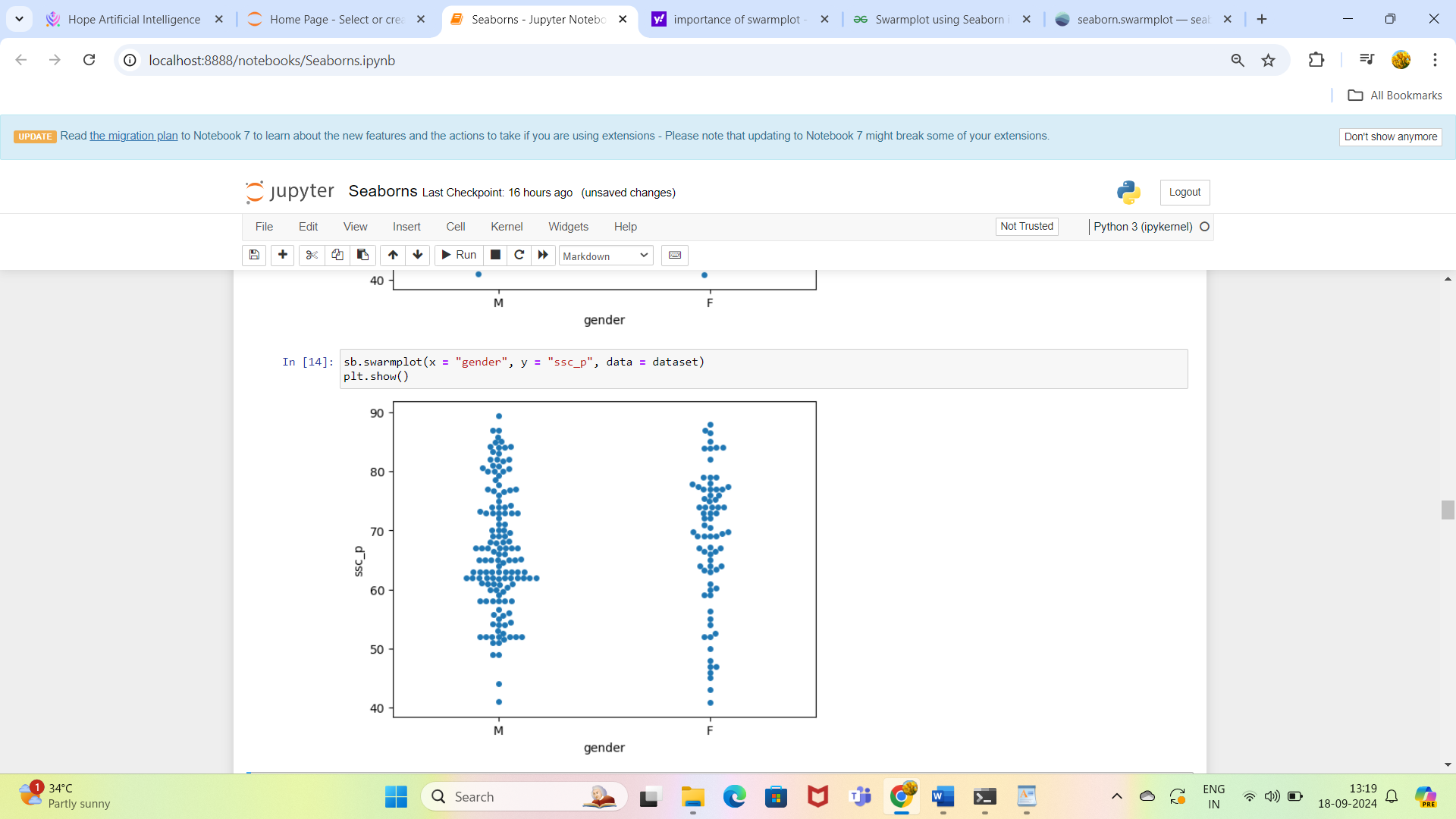
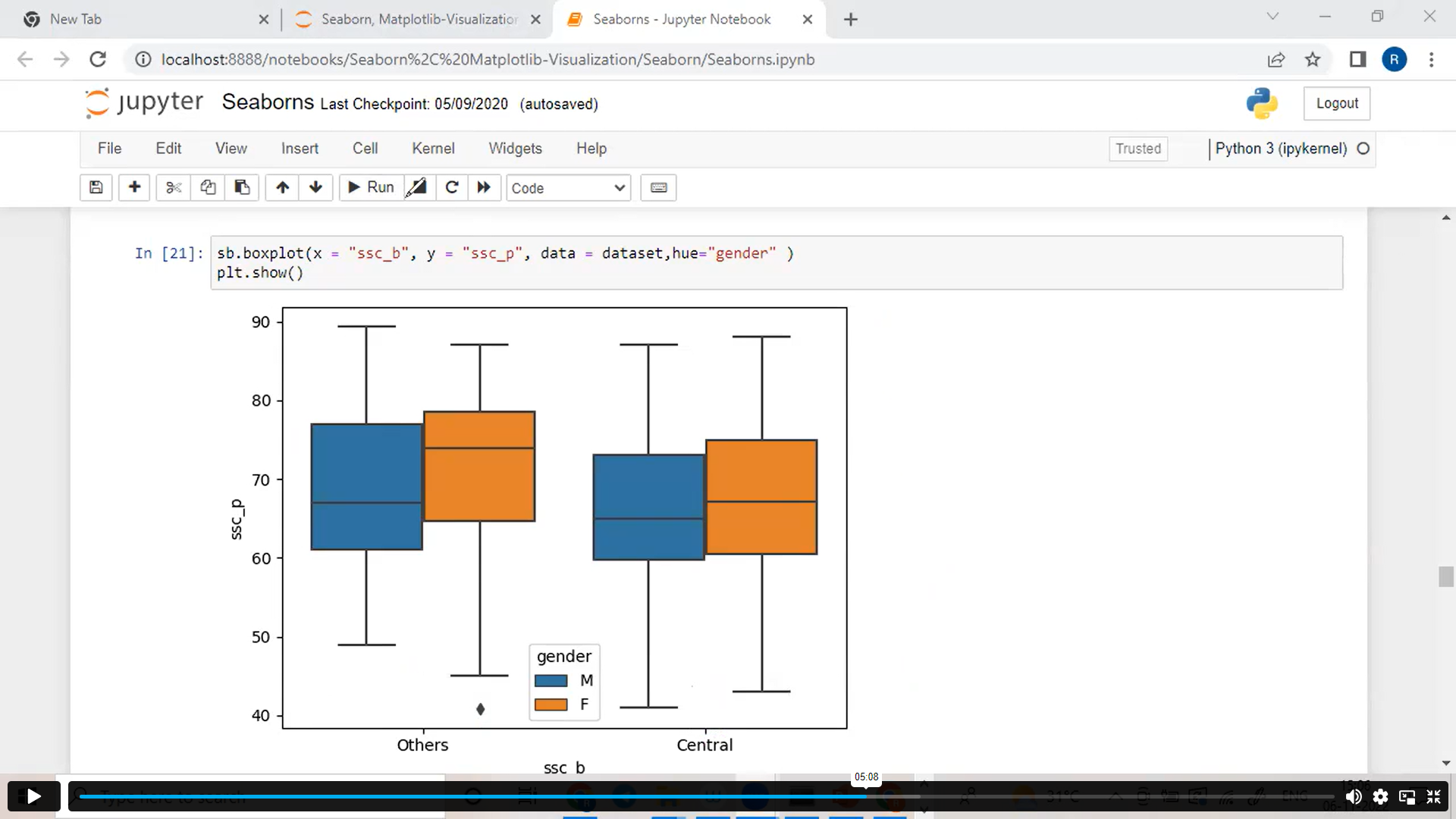
DATAVISUALIZATION -SEABORN ASSIGNMENT

# SWARMPLOT

This function is similar to stripplot(), but the points are adjusted (only along the categorical axis) so that they don’t overlap. This gives a better representation of the distribution of values, but it does not scale well to large numbers of observations. This style of plot is sometimes called a “beeswarm”.



# BOXPLOT



**BoxPlot**:

It describes about Percentile concept.

**ssc\_b =Others, Gender=Male/Female**

Others Category:

For Male ->

Initial range is starting from 50% . Q1 : 62% , Q2 : 66% , Q3 :78% ,Q4 : 89%.

For Female ->

Initial range is from 45%. Q1 :65%, Q2 :75% , Q3 :79% , Q4 : 87%.

Outlier at 40%.

So here for this Others category , we can say that compared to female students Male students has performed well.

**ssc\_b =Central, Gender=Male/Female**

Others Category:

For Male ->

Initial range is starting from 42% . Q1 : 60% , Q2 : 65% , Q3 :75% ,Q4 : 87%.

For Female ->

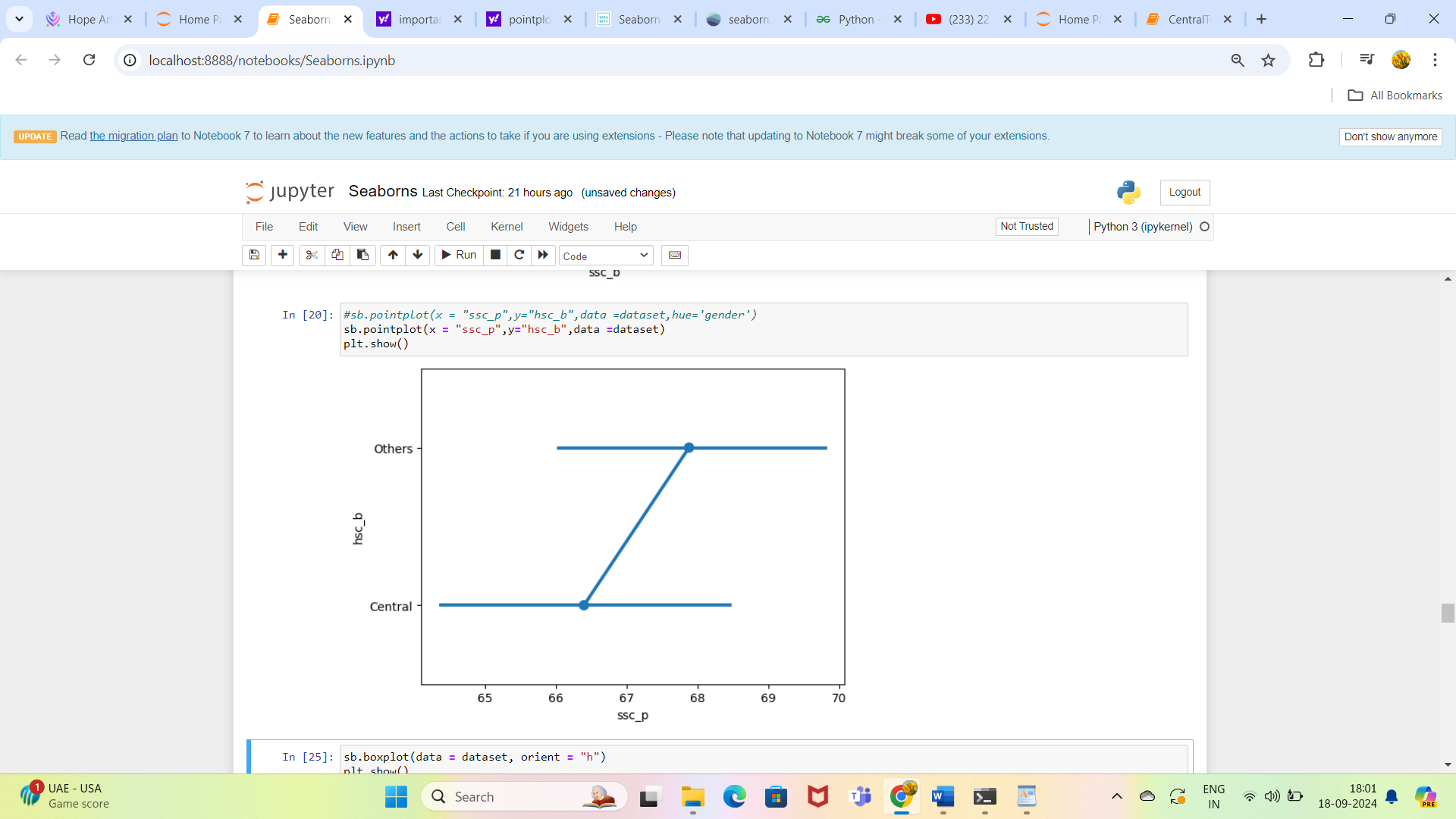
Initial range is from 44%. Q1 :62%, Q2 :68% , Q3 :77% , Q4 : 88%.

So here for this Central category , we can say that compared to Male students female students has performed well.

# POINTPLOT

Pointplot is used to calculate the centraltendency of categorical values across numerical values.

Without hue, Mean of others category is large than Central category of hsc\_b

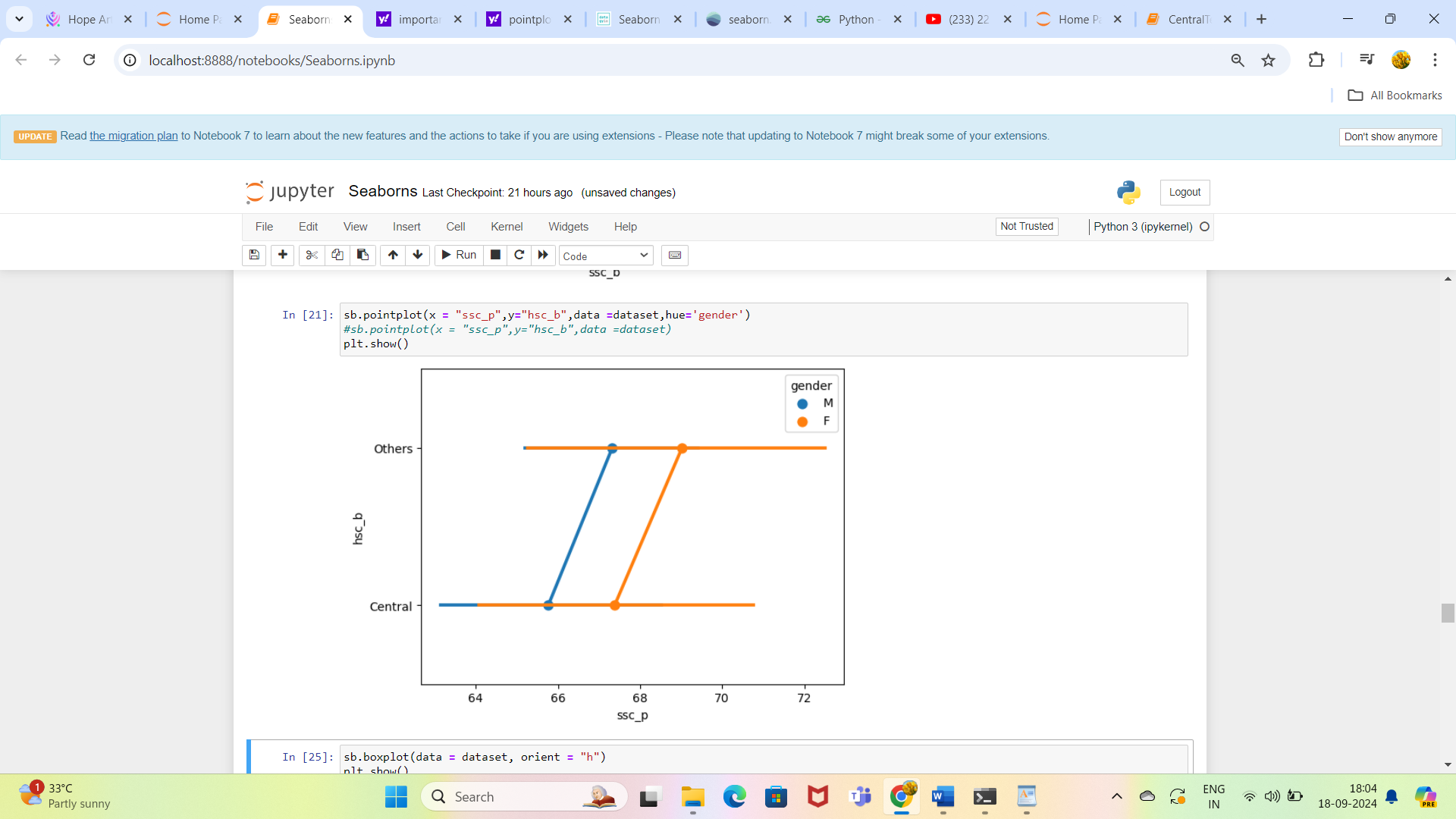


With hue based on gender

Central , Male -> Mean is smaller than Others Male

Central ,Female -> Mean is smaller than Others Female

So overall Others category has performed well.



# FACTORPLOT

The name factorplot is deprecated and now renamed as catplot.

Here we are loading the inbuilt dataset exercise.

Factorplot is used to describe the relationship between a numerical and one or more categorical variables using one of several visual representations

Here in this graph we are taking 2 categorical columns diet and kind column against the numerical column pulse.

**kind**

The kind of plot to draw, corresponds to the name of a categorical axes-level plotting function. Options are: “strip”, “swarm”, “box”, “violin”, “boxen”, “point”, “bar”, or “count”.

Here in this graph we are giving kind=**violin for visual representation.**

**row, col**

*names of variables in data or vector data*

Categorical variables that will determine the faceting of the grid.

Here we have visually represented by column wise for the categorical column **diet.**

So from the graph we can say that

Pulse rate is more for running activity compared to walking/rest mode. This is for both no fat or low fat diet.

Pulse rate is more for 30mins duration when with no fat diet where as it is huge for 15 mins time with low fat diet.

