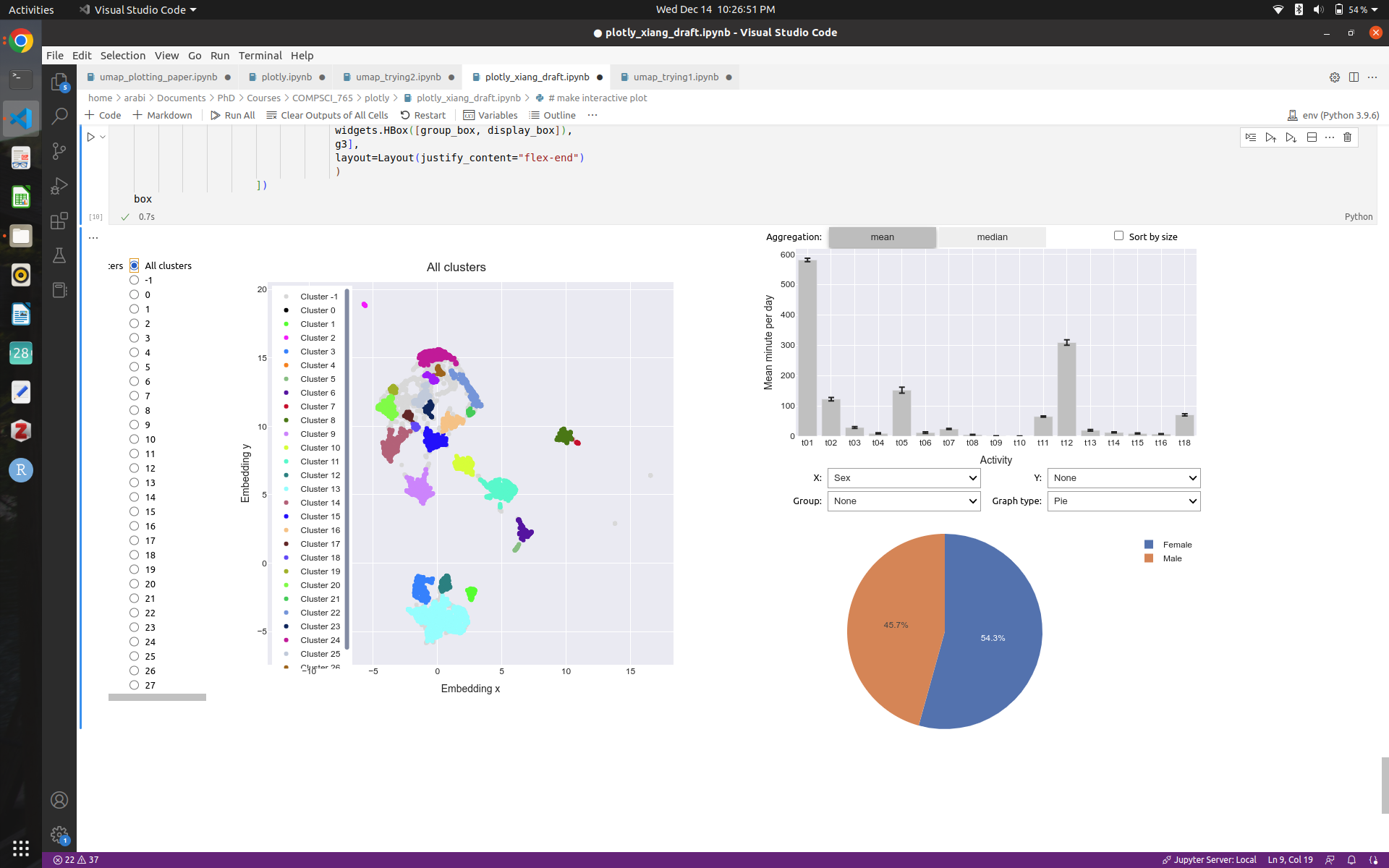
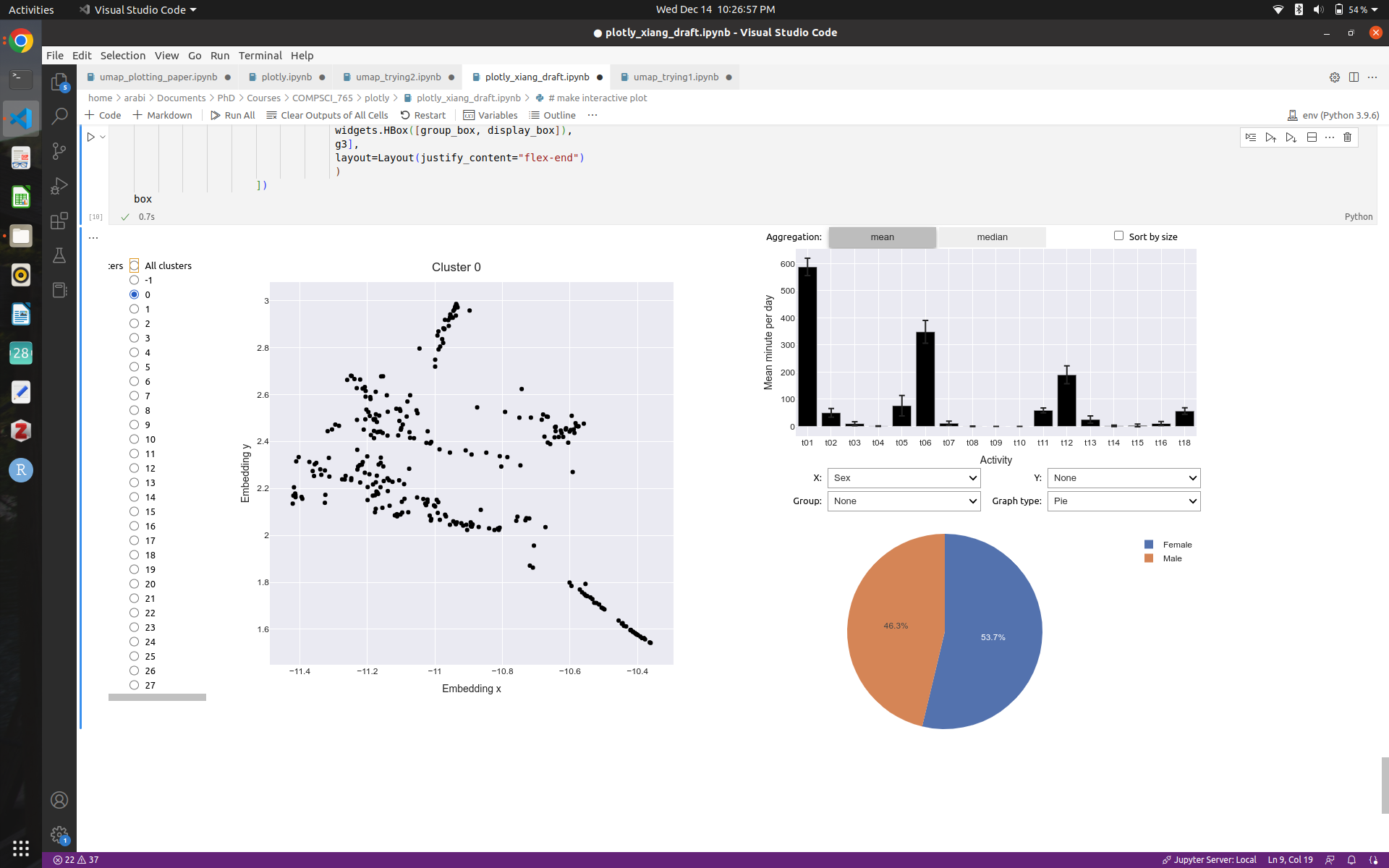
Hello, this is Xiang Zheng and Salsabil Arabi. Our project belongs to Theme 2, maps of clusters. In our project, we used an embedding approach, UMAP, to reduce the high dimensional time usage data into 2D embeddings, and then used HDBSCAN package to cluster the respondents into 28 clusters, except for the Outliers. We made a 2D interactive map to explore the time usage patterns and multi-variate relationships within each cluster and check if the clustering result is good.

1. The default layout shows the map of the data points where each individual in a dot. People within the same cluster are color-encoded. Their UMAP embeddings are encoded by positions along the x- and y-axes. At the right, we can see bar plot of the time usage of all the population and a default chart of demographic information.



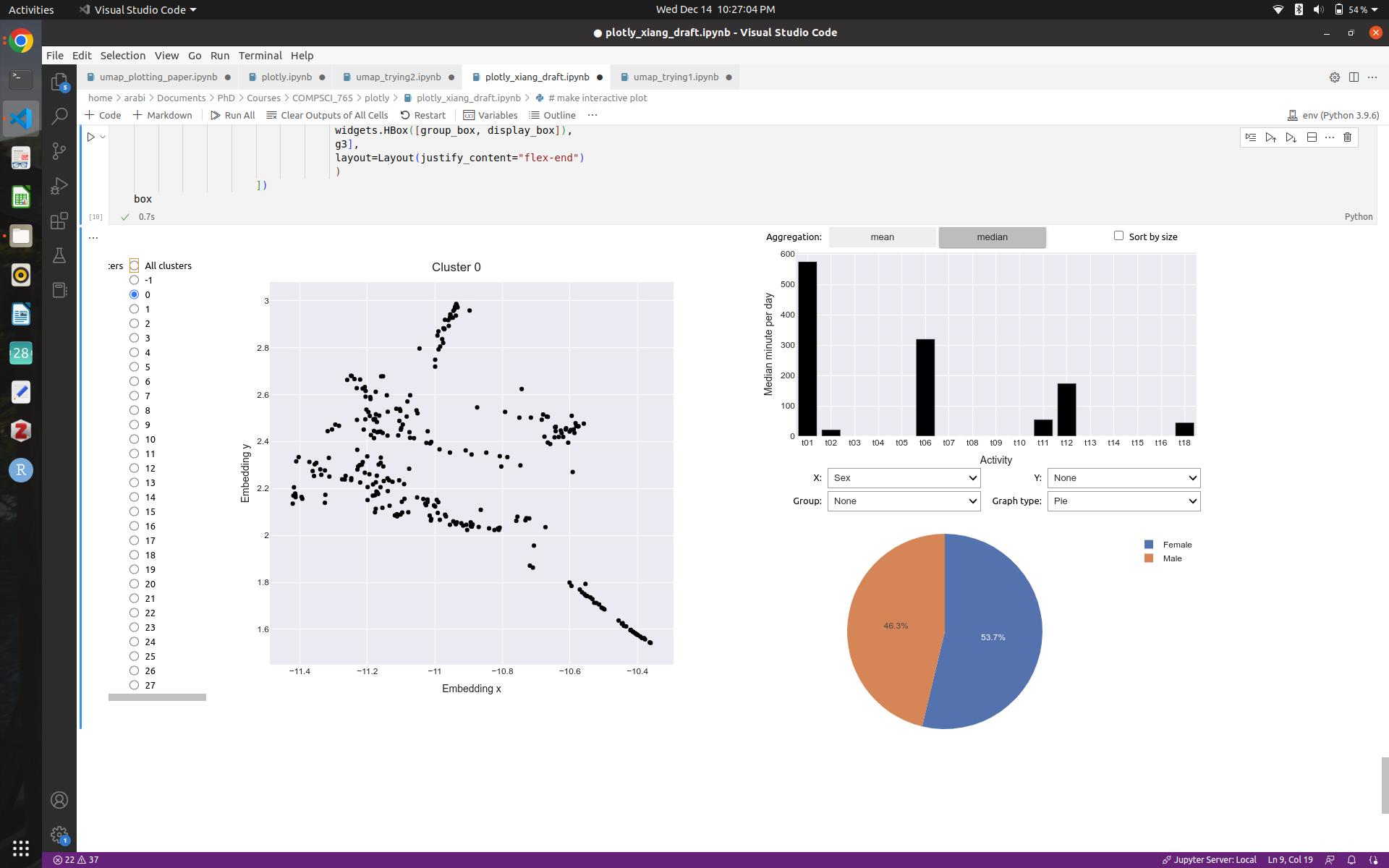
1. We can select any particular cluster number from the left bar or click any dot in the scatter plot to see the time usage and relevant information particular to the cluster. When we click on a particular cluster, all the other graphs are changed, and they show the graphs only based on the selected cluster. Due to space limitations, we did not change the code of the time usage in the bar chart in upper right. The time encodings are in Table 1.



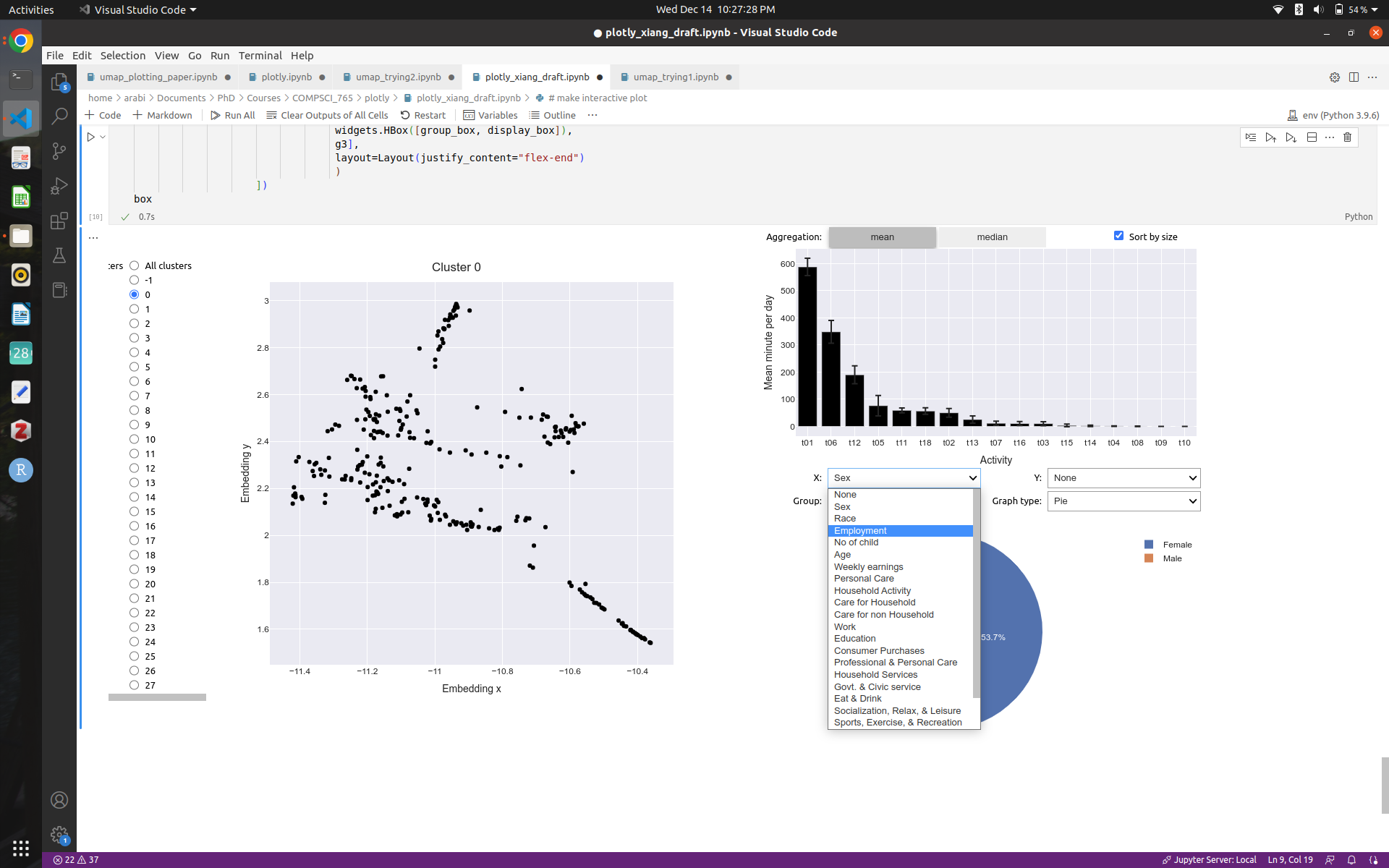
|  |  |  |  |
| --- | --- | --- | --- |
| Personal Care (t01) | Household Activities (t02) | Caring for household members (t03) | Caring for non household members (t04) |
| Work and Work Related (t05) | Education (t06) | Consumer Purchases (t07) | Professional and Personal care services (t08) |
| Household Services (t09) | Govt. & Civic Obligation (t10) | Eating and Drinking (t11) | Socializing, Relaxing, and Leisure (t12) |
| Sports, Exercise, and Recreation (t13) | Religious and Spiritual Activities (t14) | Volunteer Activities (t15) | Telephone Calls (t16) |

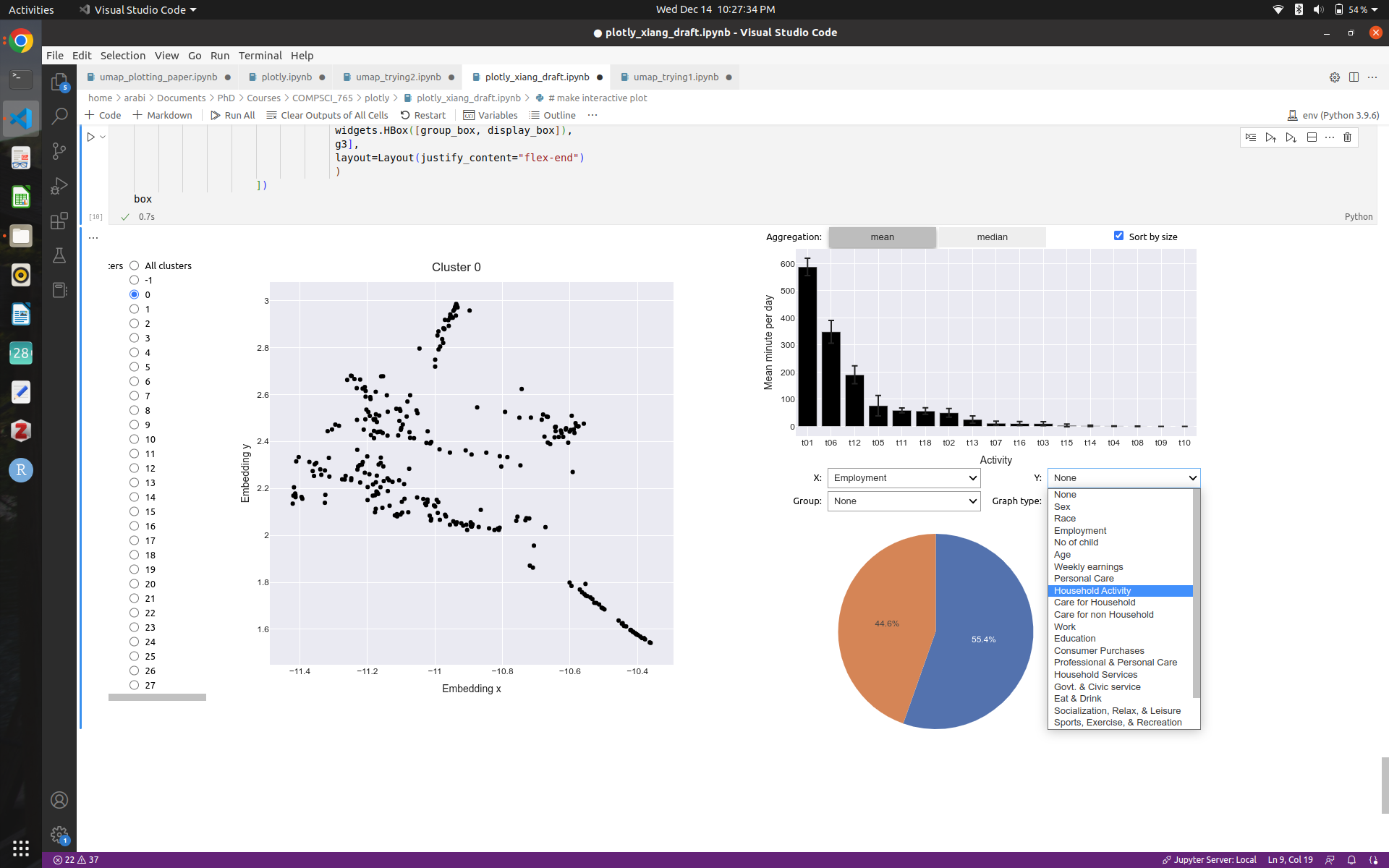
Table 1: Time usage encodings

1. From the upper right aggregation bar, we can select mean or median for the time usage of any particular cluster. By default, it shows the mean. We can see the median as mean might be skewed sometimes by outliers. We can sort the times by clicking the sort by size button.

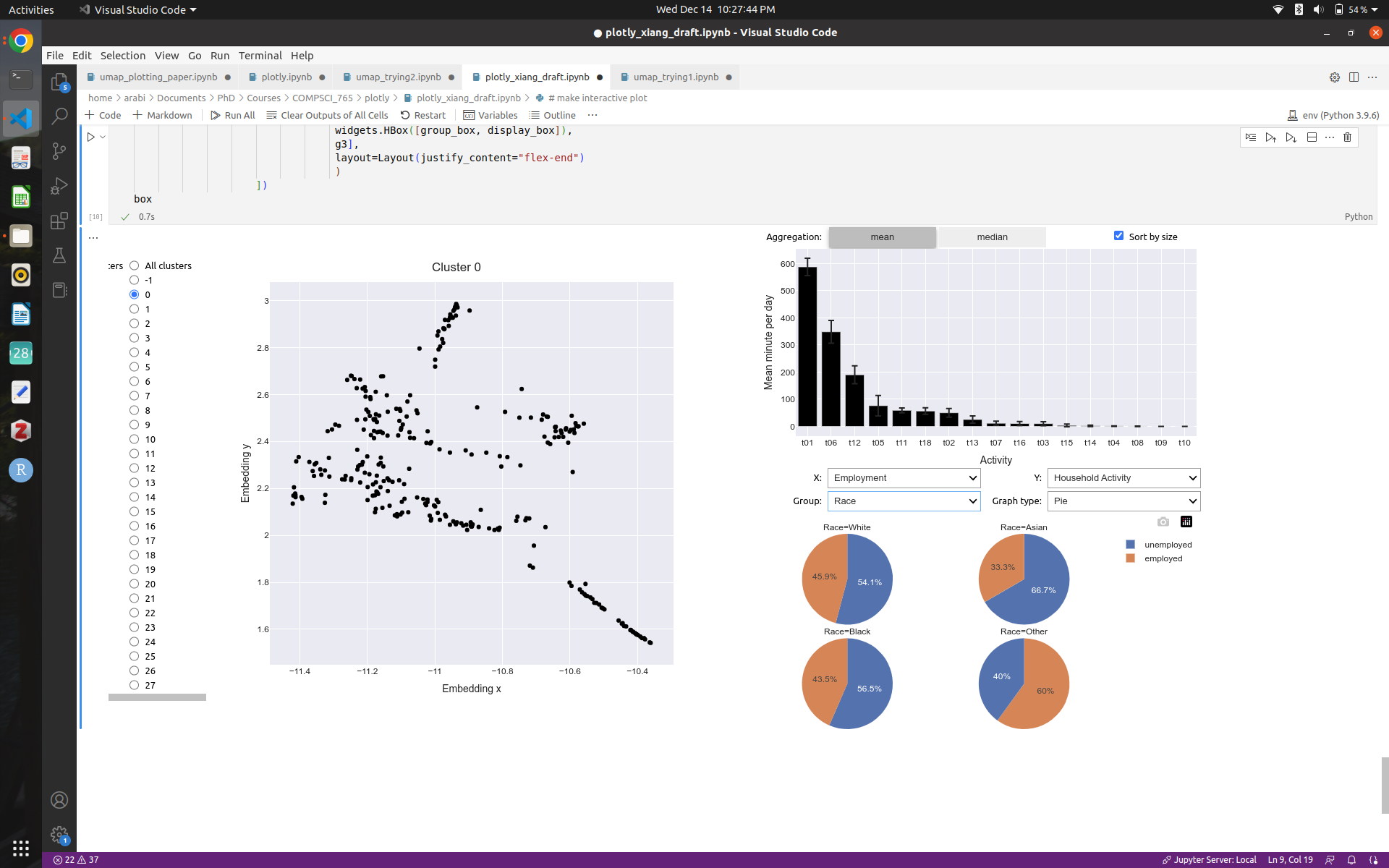


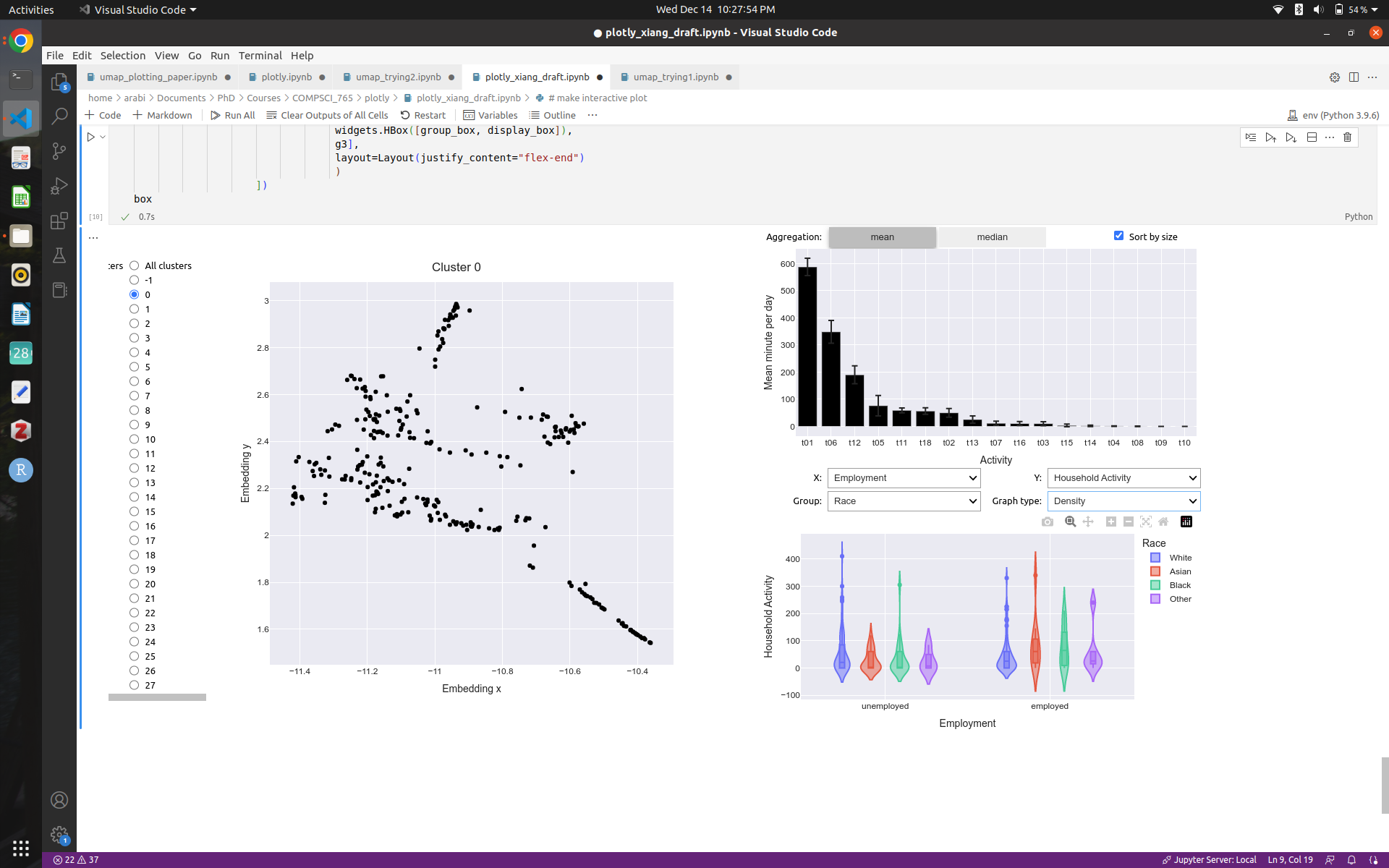
1. We can select the “x” and “y” buttons in the bottom center to see the distribution/density of the population of a particular cluster. It is a drop-down menu which lets us select from a large number of features. We can interchange “X” and “Y” as we want.

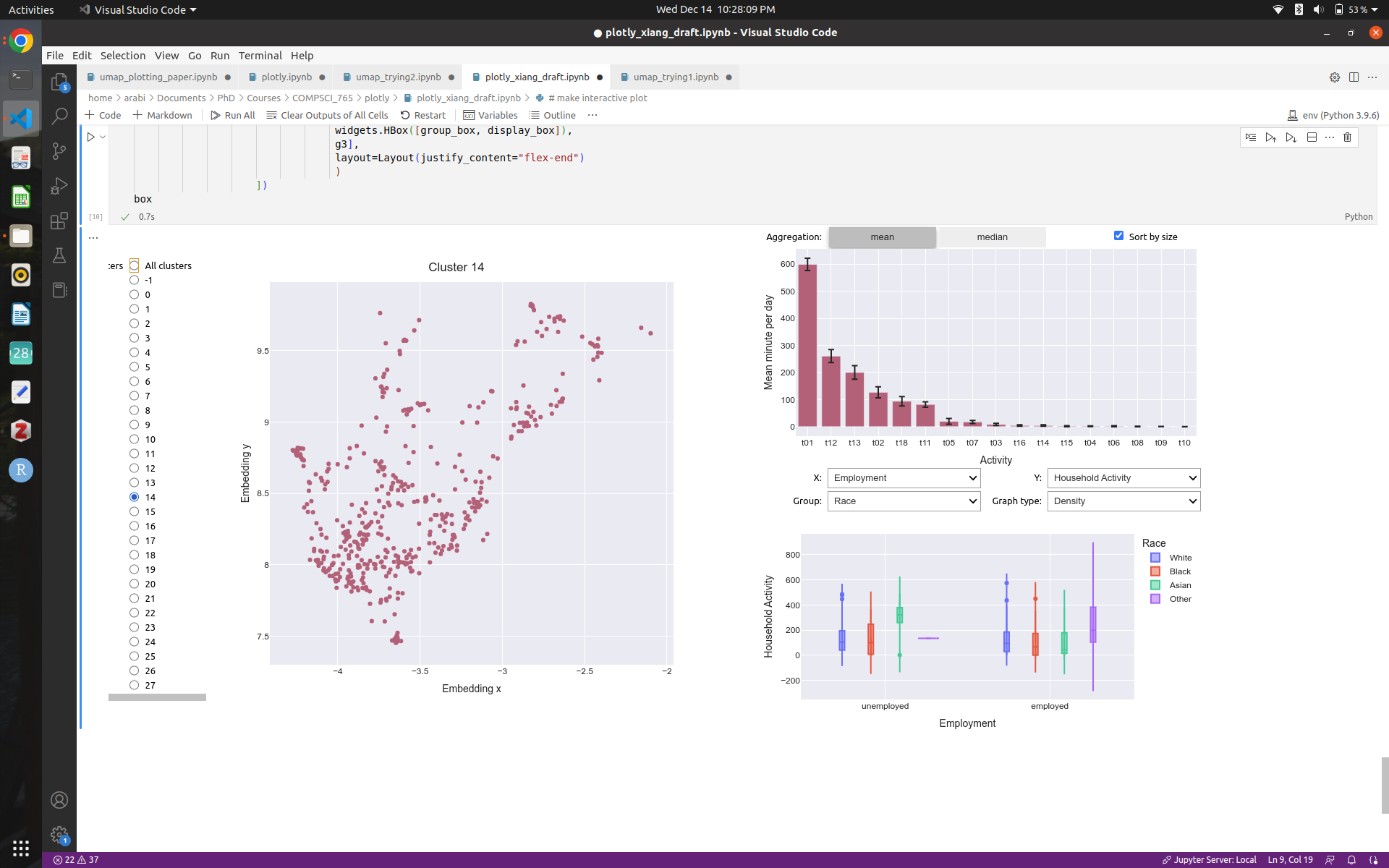




1. We can group the charts based on different categorical variables such that sex, race, and others using the “Group” button.







1. We can select different chart types from pie, density, histogram by using the “Graph type” button.

