**Data Preparation**

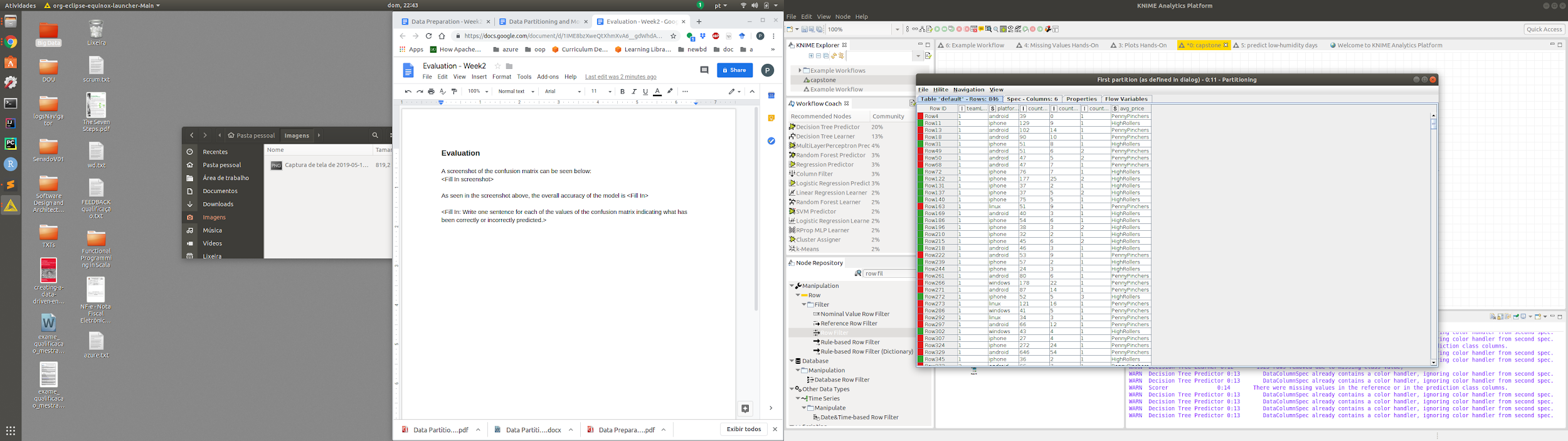
Analysis of combined\_data.csv

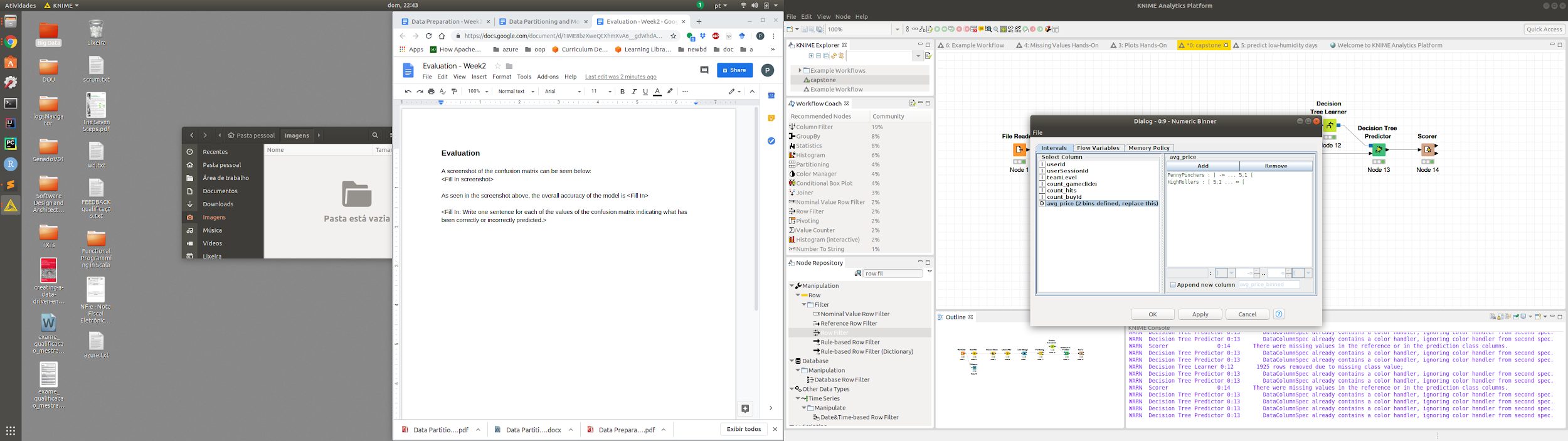
Sample Selection

|  |  |
| --- | --- |
| **Item** | **Amount** |
| # of Samples | 4619 |
| # of Samples with Purchases | 1411 |

Attribute Creation

A new categorical attribute was created to enable analysis of players as broken into 2 categories (HighRollers and PennyPinchers). A screenshot of the attribute follows:





Two groups based on average price were created, splitting the dataset into purchase above and

below $5.00, named as HighRollers and PennyPinchers.

The creation of this new categorical attribute was necessary because to be able to train a

categorization.

Attribute Selection

The following attributes were filtered from the dataset for the following reasons:

|  |  |
| --- | --- |
| **Attribute** | **Rationale for Filtering** |
| userId | Index is not valuable for the classification |
| userSessionId | Index is not valuable for the classification |
| <Optional Fill in> | <Optional Fill in 1-3 sentences> |
| <Optional Fill in> | <Optional Fill in 1-3 sentences> |