

Data Preparation

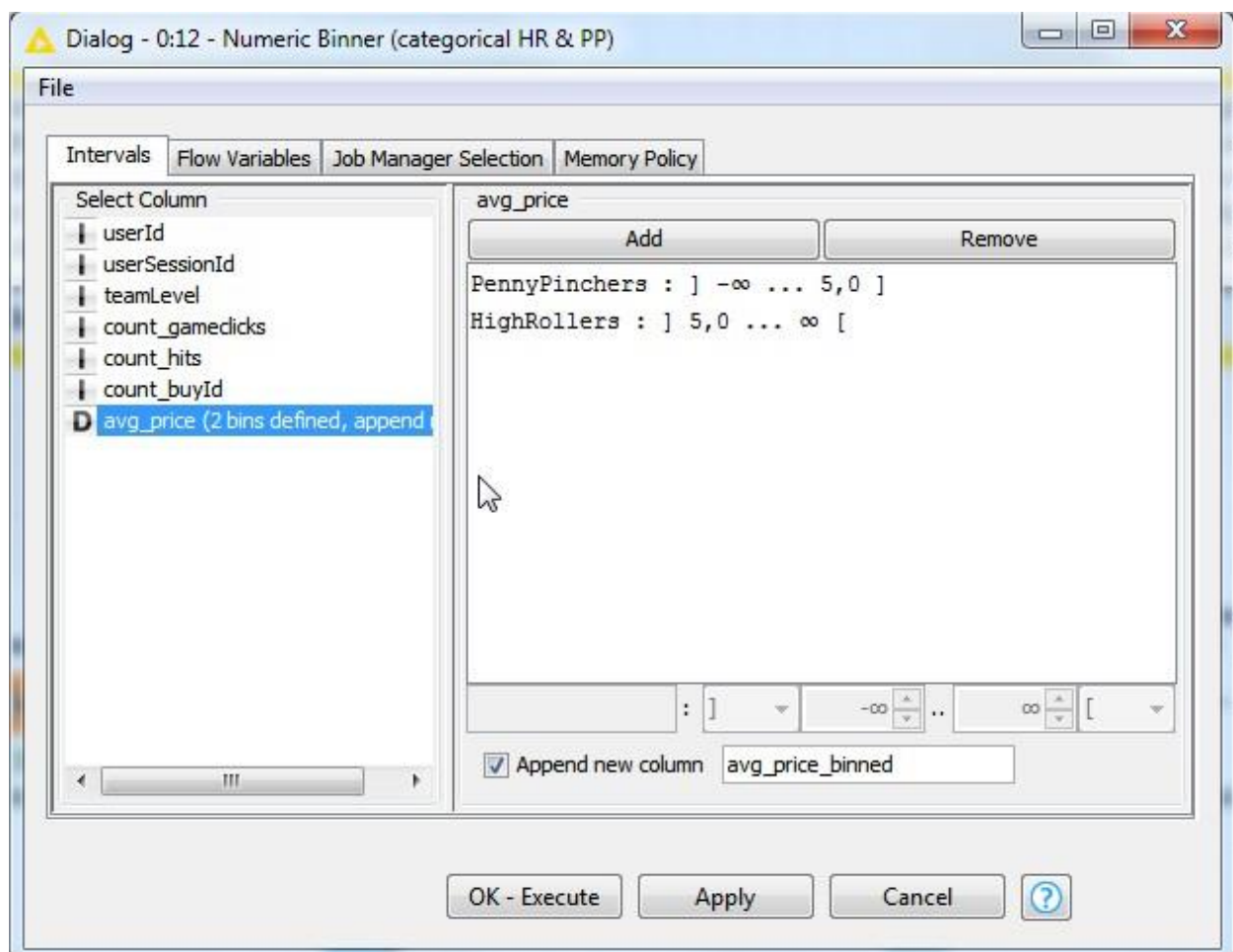
Analysis of combined_data.csv

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:



The goal of this new attribute will separate the gamers into two categories : the first ones, PeenyPinchers, who usually spent 5\$ or less and the second ones, HighRollers, who spent usually at least 5\$.

The creation of this new categorical attribute was necessary because predicting which users will spent more or less money in the in-app purchases is a valuable information for Eglence Inc.

Attribute Selection

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

Attribute	Rationale for Filtering
Avg-price	The attribute doesn't depend directly of the user
userId	Information originating from our server
Usersessionid	Information originating from our server