

Data Preparation

Analysis of combined_data.csv

Sample Selection

Item	Amount
# of Samples	4619
# of Samples with Purchases	1411

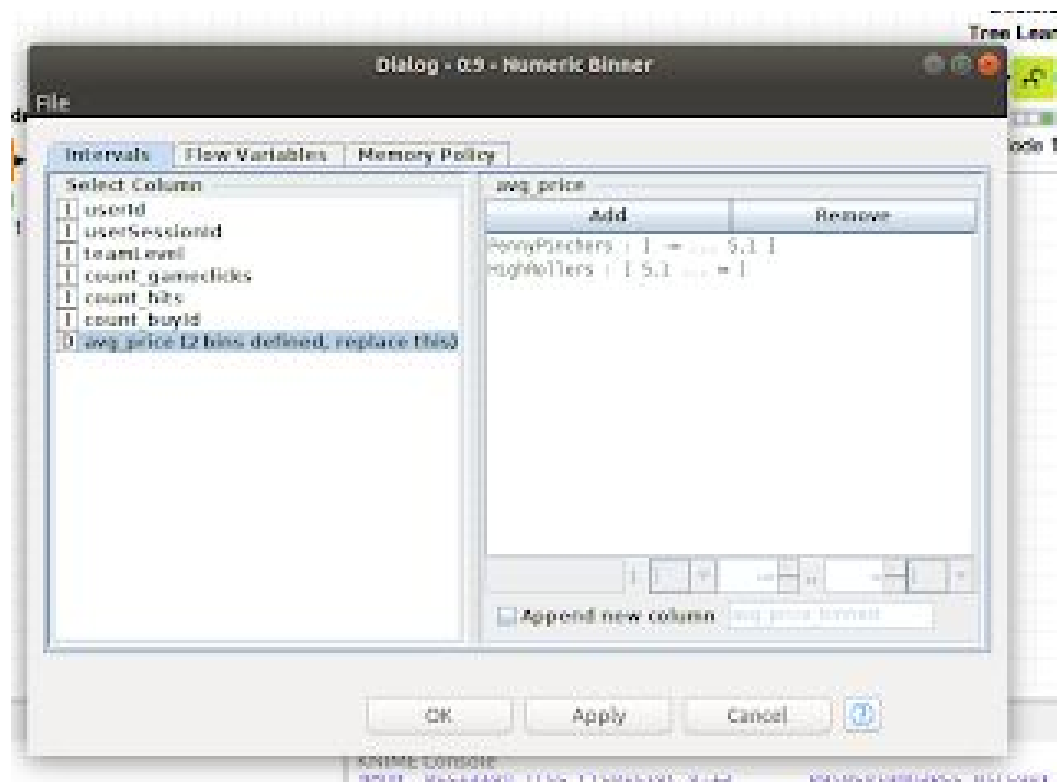
Attribute Creation

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 screenshot shows a data table with the following columns: Row ID, team, platform, count, count, count, and avg_price. The data is categorized into two groups: HighRollers and PennyPinchers. The table is titled 'Table: default - Rows: 846' and 'Spec - Columns: 6'. The 'Platform' column lists various operating systems like android, ios, linux, windows, and macos. The 'count' column shows the frequency of each platform. The 'avg_price' column shows the average price for each platform. The 'team' column lists various teams like PennyPinchers, HighRollers, and others. The 'Row ID' column lists row numbers from 1 to 846.

Row ID	team	platform	count	count	count	avg_price
Row1	1	android	39	0	1	PennyPinchers
Row11	1	ios	129	9	1	HighRollers
Row13	1	android	102	14	1	PennyPinchers
Row18	1	android	90	10	1	PennyPinchers
Row21	1	ios	51	8	1	HighRollers
Row49	1	android	51	6	2	PennyPinchers
Row55	1	android	47	5	2	PennyPinchers
Row68	1	android	47	7	1	PennyPinchers
Row72	1	ios	36	7	1	HighRollers
Row122	1	ios	177	25	2	HighRollers
Row131	1	ios	37	2	1	HighRollers
Row137	1	ios	37	5	2	HighRollers
Row140	1	ios	75	5	1	HighRollers
Row163	1	linux	51	9	1	PennyPinchers
Row169	1	android	40	3	1	HighRollers
Row186	1	ios	54	6	1	HighRollers
Row196	1	ios	38	3	2	HighRollers
Row210	1	ios	32	2	1	HighRollers
Row215	1	ios	45	6	2	HighRollers
Row218	1	android	46	3	1	HighRollers
Row222	1	android	53	8	1	PennyPinchers
Row239	1	ios	57	2	1	HighRollers
Row244	1	ios	26	3	1	HighRollers
Row261	1	android	60	6	1	PennyPinchers
Row266	1	windows	178	22	1	PennyPinchers
Row271	1	android	87	14	1	PennyPinchers
Row272	1	ios	52	5	3	HighRollers
Row273	1	linux	121	16	1	PennyPinchers
Row286	1	windows	41	5	1	PennyPinchers
Row292	1	linux	34	3	1	PennyPinchers
Row297	1	android	66	12	1	PennyPinchers
Row302	1	windows	43	4	1	HighRollers
Row307	1	ios	27	4	1	PennyPinchers
Row324	1	ios	272	24	1	PennyPinchers
Row329	1	android	646	54	1	PennyPinchers
Row345	1	ios	36	2	1	HighRollers



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>

