Suraj Thyagaraj

CWRU Data Analytics Boot Camp

Spring 2020

Module 02-Python-Pandas | HW4

Topics Covered: pandas, jupyter notebook, dataframes, groupby, binning, mapping

Background

Use Python programming in a data munging challenge. In Heroes of Pymoli, the goal is to create a Jupyter Notebook to analyze the sales data of a gaming company. Using dataframes and data querying techniques in Pandas, create summary reports of client demographics, and sales reports by categories.

REPORT:

The Jupyter Notebook codes is named: **HeroesOfPymoli\_ST.ipynb**

Steps Followed:

1. Add the necessary dependencies and read in the given sales data file into a pandas dataframe.

2. For reporting summary stats of user demographics, comb through the data frame and compile the necessary metrics.

3. For reporting sales data classified by variables such as user age or game title, use groupby() to create sub-groups of the complete dataset, that are ordered by the variables under consideration.

4. After creating these sub-groups, compute the necessary metrics and apply proper formatting using mapping.

Observed Trends:

1. A vast majority of users are male (84%).

2. About half of the total users fall in the 20-24 age group (~45%).

3. Our most popular and most profitable item is Item ID:178, ‘Oathbreaker, Last Hope of the Breaking Storm’.