**Extract, Transform, Load**

Project 2 Final Report

Group Adam Bilski

Alan Riveros

Aja Tashjian

**Project Scope:**

To compare stats of basketball players with their current annual salary.

**Extract**

We scraped data from basketball-reference.com for the basketball players stats as well as downloaded .csv files on basketball salaries and the highest paid athletes from data.world and basketball-reference.com.

**Transform**

We loaded our data into pandas in order to capture only the data wanted as well as prep for adding to Postgres. Many columns in the Pandas data frames required naming and formatting changes and some required parsing out data in columns in order to keep a specific part of the data. Any columns with extraneous data were dropped.

**Load**

Once data was cleaned it was loaded into Postgres using Pandas into three tables:

* Players (basketball player names and their current salaries)
* Top\_35 (basketball player names including their pay, winnings, endorsements and marital status)
* Stats (basketball player names and their current stats)

This allows for cross-referencing a basketball player’s stats with their earnings (or perhaps marital status).