**ETL Project Report**

*Erik Sjostrom & Zachary Pitts*

* Extract

The sources of data that we extracted were NBA players stats from the 2012-2015 seasons. Player’s height, points, rebounds, and win shares. We discovered the datasets from Kaggle and the Google datasets search engine. We chose to extract the stats from the NBA championship teams from their respective years. We did this to try to answer the question if these stats correlate to winning a championship. We formatted these datasets on pandas into CSV format.

* Transform

The type of transformations that were necessary for these datasets were filtering the data to the 3 years that we used. Also, transforming players heights from centimeters to inches, while cleaning to only two decimal points. Transforming the data into tables in pandas was another step we took.

* Load

After loading the raw datasets in CSV format to pandas and cleaning. We transform those new tables into CSV format and uploaded to pgAdmin4. Created a new database named ETL-Project and placed our tables into this database.