

# Predicting Fantasy Football Performances

John O'Hollaren (jpo4@duke.edu)  
Department of Electrical and Computer Engineering  
Duke University



## Motivation

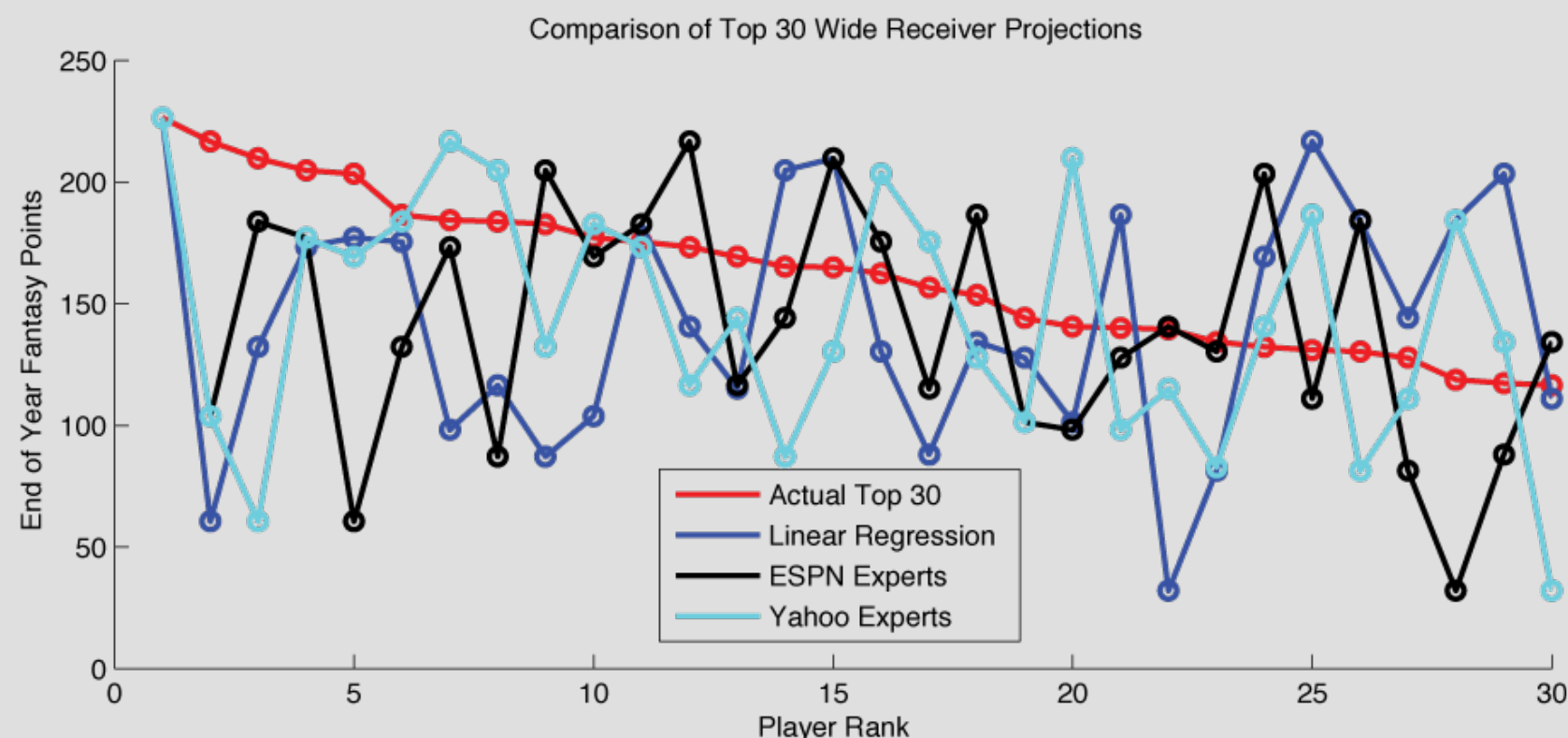


Figure 1: Total end-of-year fantasy points scored by ESPN / Yahoo's preseason top 30, compared to the actual top 30 and the featured reduced linear regression algorithm. This shows that ranking is very difficult.

### Why is predicting fantasy football stats important?

- **Profit:** \$15 billion per year is spent on fantasy football.
- **Popularity:** According to the Fantasy Sports Trade Association, 10% of the US population plays fantasy football.
- **Difficulty:** Ranking players is difficult, even the experts are rarely correct (see Figure 1). A small edge can therefore go a long ways.

### How does fantasy football work?

Points are awarded when a player scores a touchdown or gains yards. Each owner picks their team at the beginning of the year. Correctly predicting which players will play well greatly improves your team's year-long performance. This paper focuses on an optimal strategy for predicting which NFL wide receivers will have the best fantasy football season using data available prior to the beginning of the season.

### Goal

Prediction of NFL receivers' end of season fantasy point ranking.

## Data

The following key statistics were gathered for every single active wide receiver in the NFL from 2007 until 2012.

700+ Players Per Year →

← 13 Features	Receiving Yards
	Receiving Touchdowns
	Receiving Targets
	Receiving Catches
	ESPN Preseason Ranking
	Yahoo Preseason Ranking
	Δ 1 year Receiving Yards: (Yards in year Y-1) – (Yards in year Y-2)
	Δ 1 year Receiving TDs: (TDs in year Y-1) – (TDs in year Y-2)
	Δ 2 year Receiving Yards: (Yards in year Y-1) – (Yards in year Y-3)
	Δ 2 year Receiving TDs: (TDs in year Y-1) – (TDs in year Y-3)
	Δ 3 year Receiving Yards: (Yards in year Y-1) – (Yards in year Y-4)
	Δ 3 year Receiving TDs: (TDs in year Y-1) – (TDs in year Y-4)
	Fantasy Points in Previous Year: 6 • TDs + 0.1 • Yards

### Performance Criteria

Discounted cumulative gain (DCG) will be used to reward ranking good players highly and penalize ranking good players low. A higher DCG is better.

$$DCG = rel_1 + \sum_{i=2}^p \frac{rel_i}{\log_2(i)} \quad rel_i = \frac{1}{\text{actual end of year rank}}$$

Below we can visually view our data, as well as the inconsistency of professional rankings. These will serve as our benchmarks.

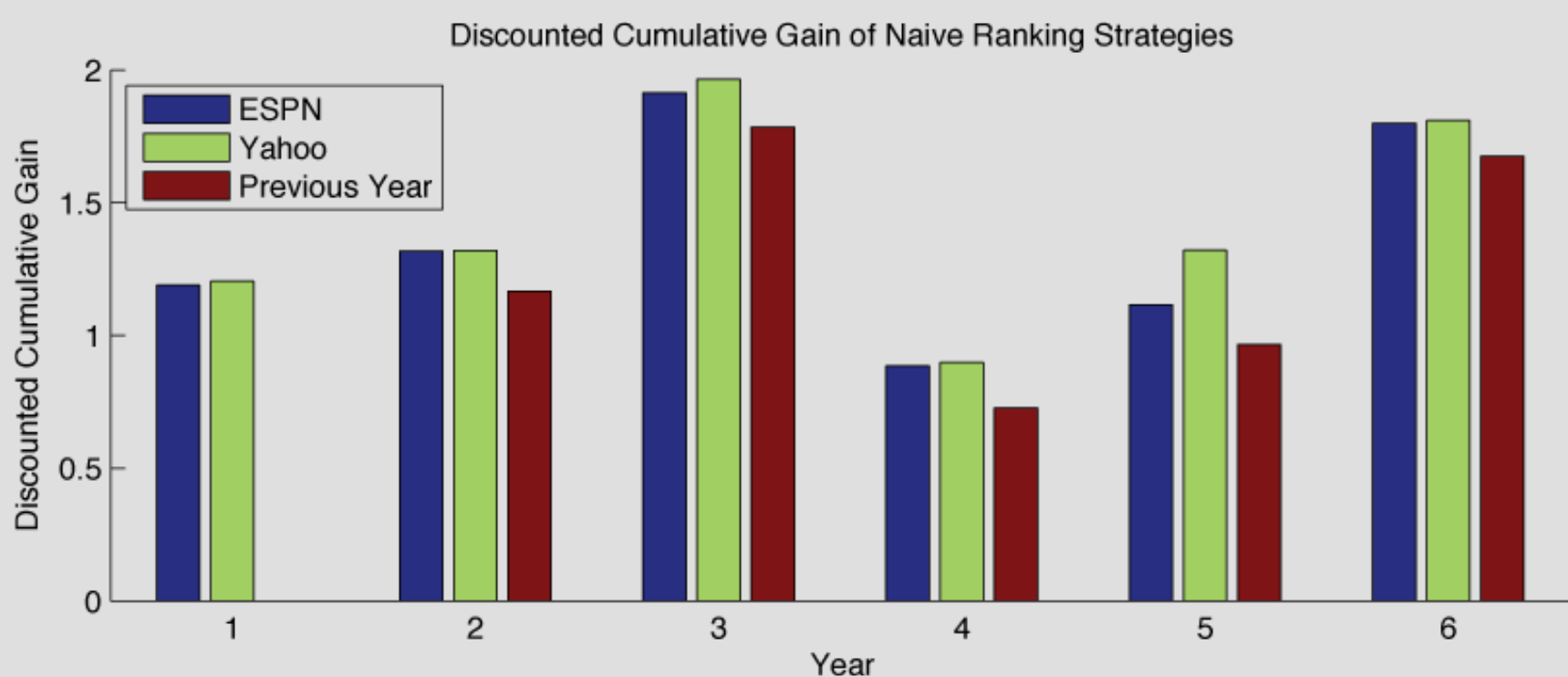


Figure 2: DCG of ESPN / Yahoo preseason rankings over the years, along with a naive method of using the previous year's final rankings as the next year's preseason rankings.

Expert	Discounted Cumulative Gain for 2012 Predictions
ESPN Expert (Matthew Berry)	1.7991
Yahoo Team of Experts	1.8100

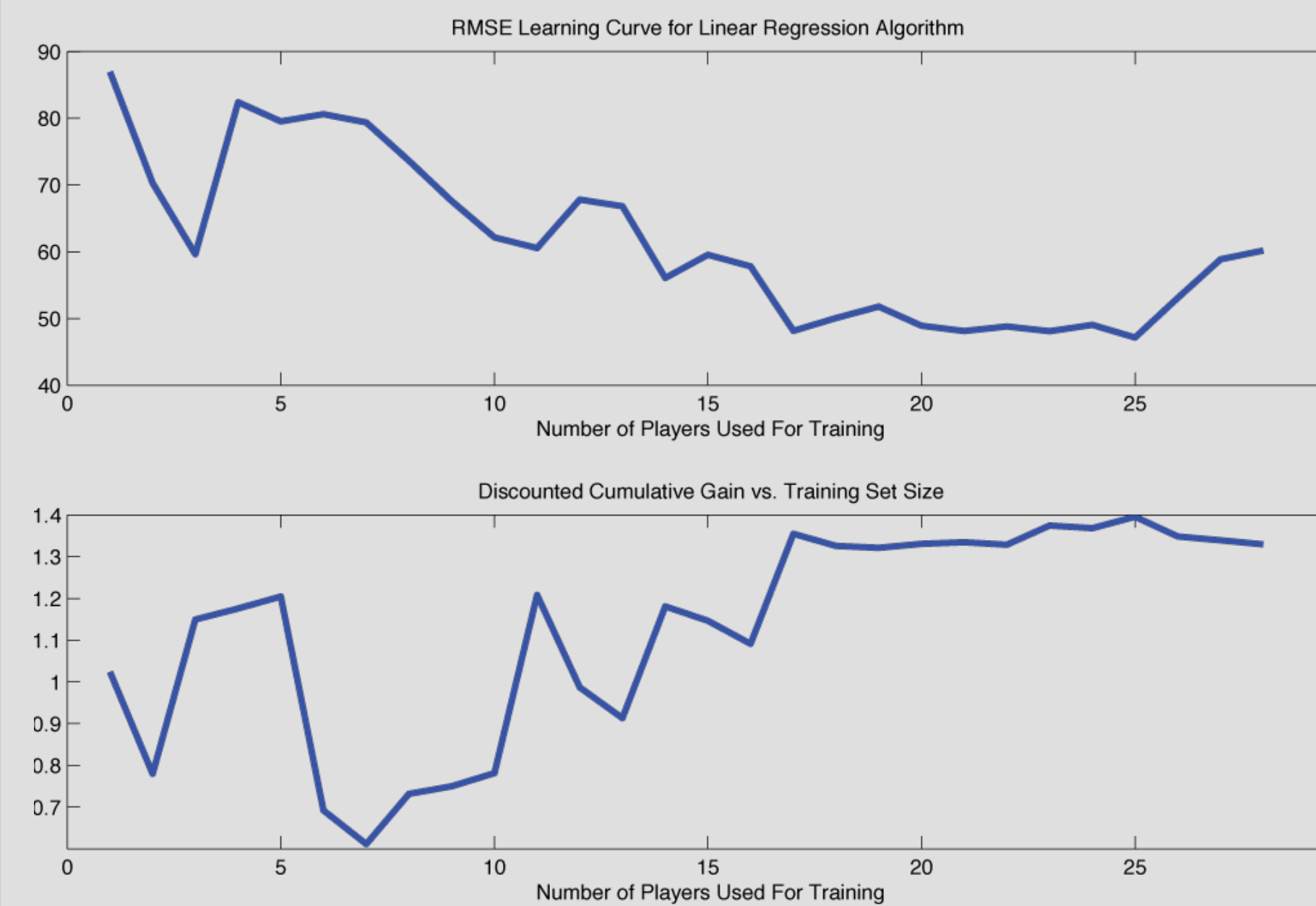
Figure 3: The numeric DCG of ESPN and Yahoo in 2012. This the benchmark used for this research.

## Methods

$$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecYds}} \ w_{\text{RecTDs}} \ w_{\text{Targets}} \ w_{\text{Catches}} \ w_{\Delta 1 \text{yrRecYds}} \ w_{\Delta 1 \text{yrRecTDs}} \\ w_{\Delta 2 \text{yrRecYds}} \ w_{\Delta 2 \text{yrRecTDs}} \ w_{\Delta 3 \text{yrRecYds}} \ w_{\Delta 3 \text{yrRecTDs}} \ w_{\text{ESPN}} \ w_{\text{Yahoo}}]$$

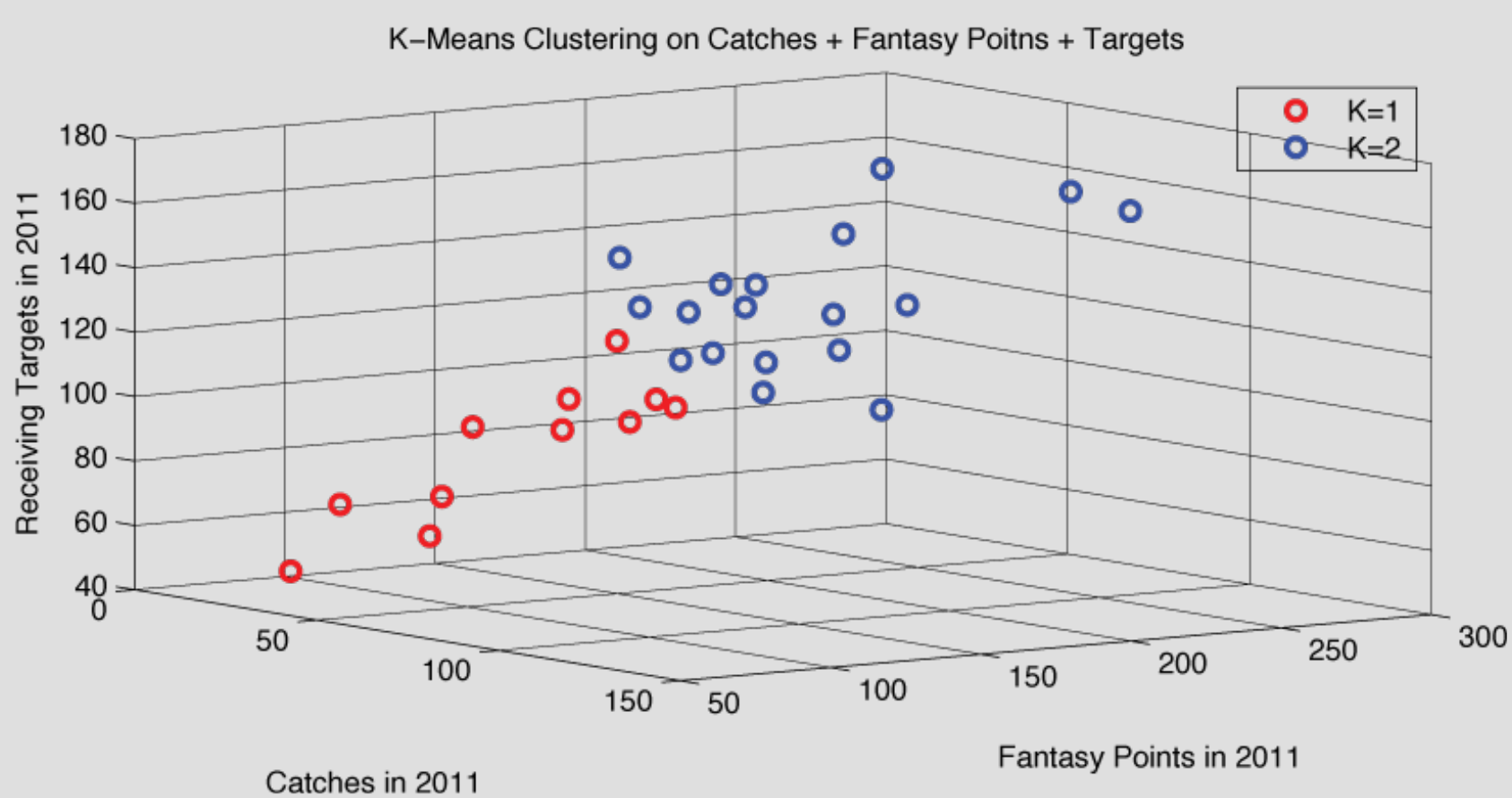
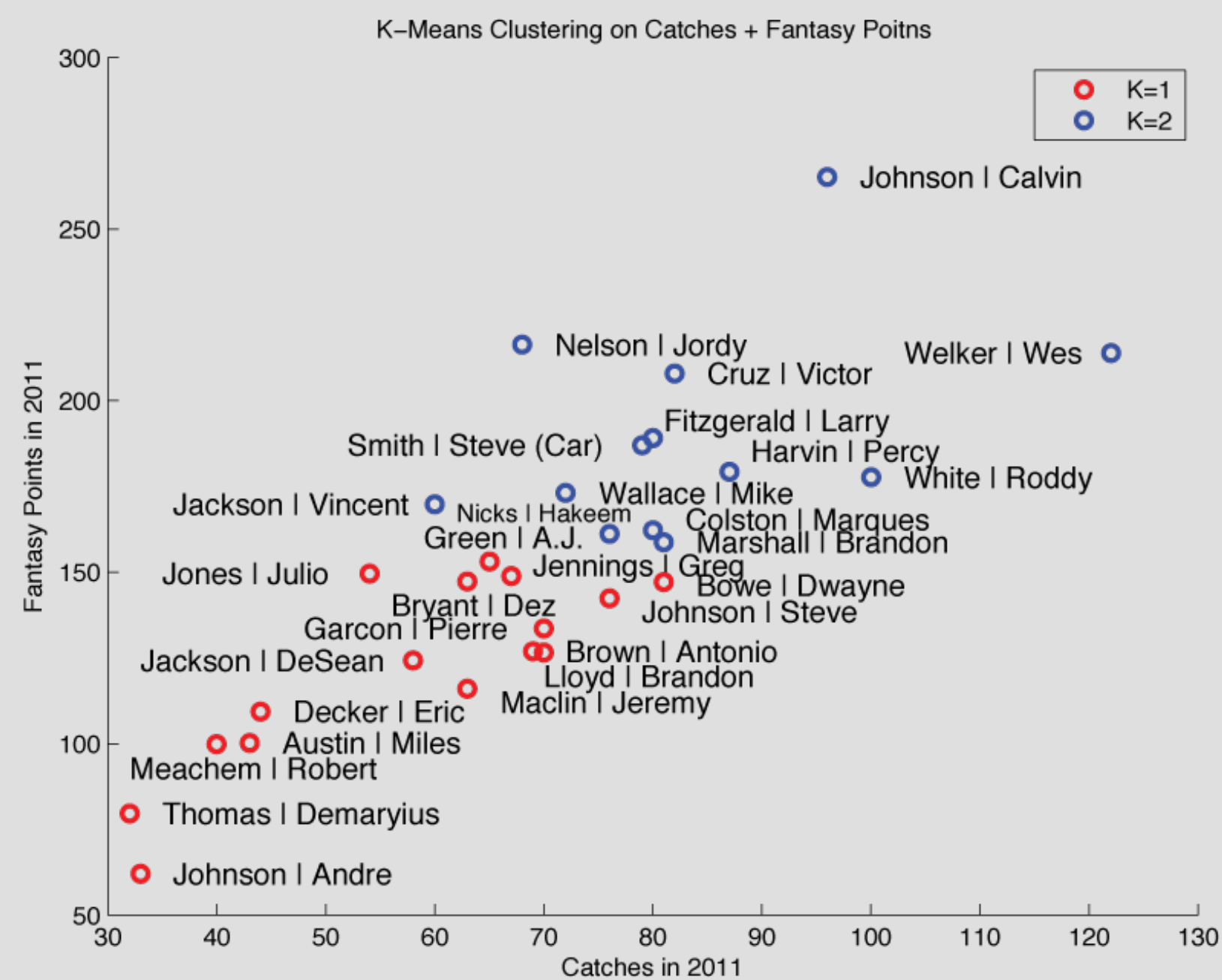
$$\beta = [w_1 \dots w_L]$$

## Linear Regression



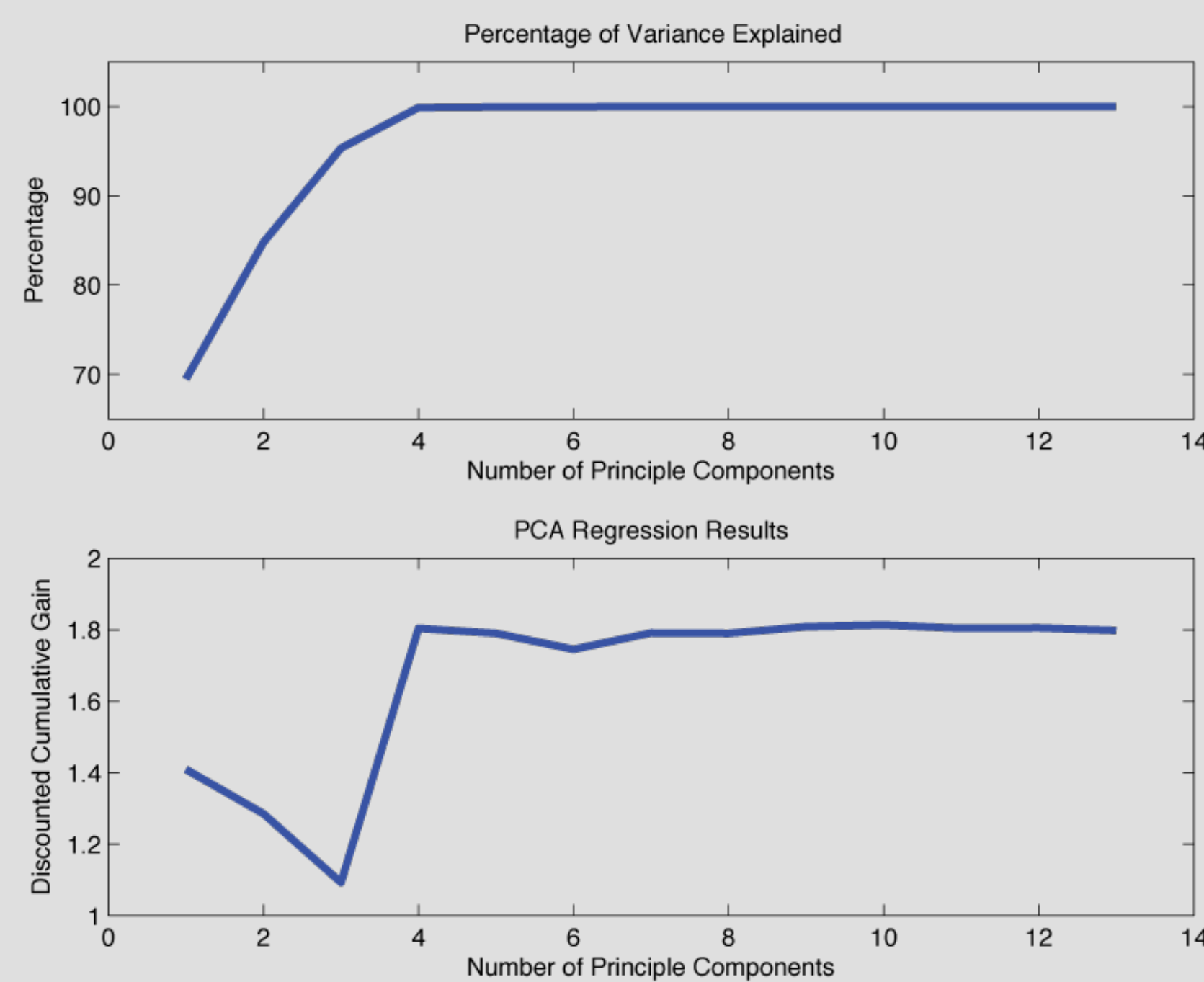
Method	$\beta$	Discounted Cumulative Gain
13 Features	$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecYds}} \ w_{\text{RecTDs}} \ w_{\text{Targets}} \ w_{\text{Catches}} \ w_{\Delta 1 \text{yrRecYds}} \ w_{\Delta 1 \text{yrRecTDs}} \\ w_{\Delta 2 \text{yrRecYds}} \ w_{\Delta 2 \text{yrRecTDs}} \ w_{\Delta 3 \text{yrRecYds}} \ w_{\Delta 3 \text{yrRecTDs}} \ w_{\text{ESPN}} \ w_{\text{Yahoo}}]$ $\beta = [0.17, 0.059, 12.91, -0.25, 0.18, -0.34, -4.82, -0.031, 0.61, 0.044, -3.36, -1.05, 1.03]$	1.7179
7 Features	$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecYds}} \ w_{\text{RecTDs}} \ w_{\Delta 1 \text{yrRecYds}} \ w_{\Delta 1 \text{yrRecTDs}} \ w_{\text{ESPN}} \ w_{\text{Yahoo}}]$ $\beta = [-0.18, 0.12, 11.10, -0.056, -3.35, -2.27, 1.67]$	1.7722
6 Features	$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecYds}} \ w_{\text{RecTDs}} \ w_{\Delta 1 \text{yrRecYds}} \ w_{\Delta 1 \text{yrRecTDs}} \ w_{\text{ESPN}}]$ $\beta = [0.10, 0.079, 10.23, -0.06, -3.09, -0.403]$	1.7621
3 Features	$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecYds}} \ w_{\text{RecTDs}}]$ $\beta = [1.72, -0.061, -5.08]$	1.7806
2 Features	$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecTDs}}]$ $\beta = [1.15, -2.031]$	1.7896
2 Features	$\beta = [w_{\text{FantasyPts}} \ w_{\text{RecYds}}]$ $\beta = [0.92, 0.019]$	1.7863

## K-Means Mixture Model



# Mixtures	K Means Features	DCG (6 Features)	DCG (2 Features)
2	Catches Fantasy Points	1.7601	1.8063
2	Fantasy Points	1.7601	1.7925
3	Fantasy Points	1.3621	1.7880
2	Receiving Targets	1.2053	1.7877
2	Catches Fantasy Points Targets	1.1887	1.7863
2	Catches	1.1811	1.7814
2	Touchdowns	1.1107	1.7747
2	Receiving Yards	1.0857	1.4428

## PCA Regression



Number of Principle Components	Discounted Cumulative Gain
4	1.8040
10	1.8132

## Conclusion

Show ESPN, Yahoo, My best, Actual. Then color code it

## References