

THE BAKER'S DOZEN ANALYTICS PRESENTS

FANTASY FOOTBALL: CONSISTENT OR ERRATIC?

BIG IDEA

- ▶ Create model to predict fantasy points that a quarterback will score in a game
- ▶ Features are averages of different statistics prior to game

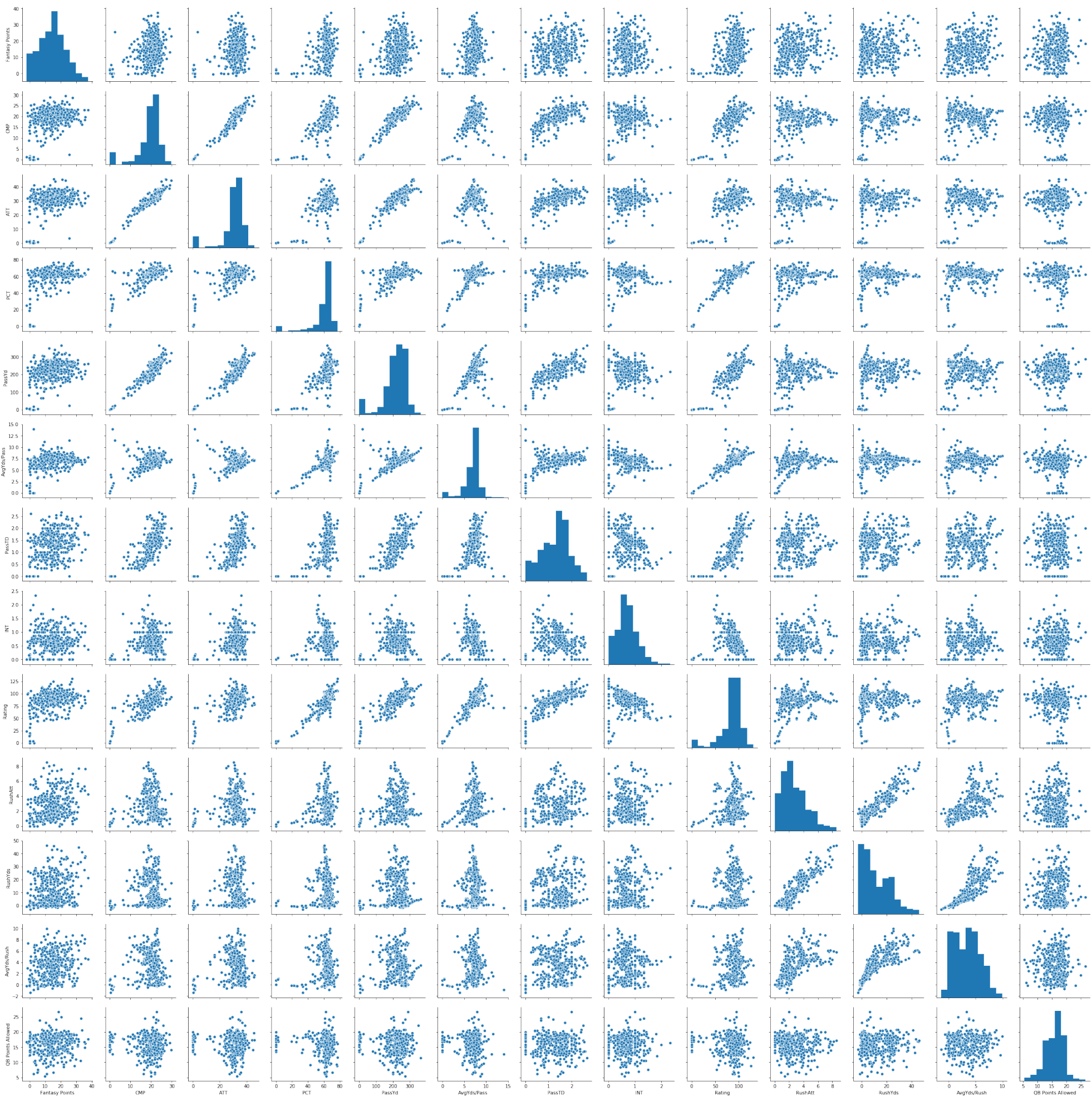
FEATURES

- ▶ Scraped QB data per game
- ▶ Scraped defensive data per game
- ▶ Final clean data:

Fantasy Points	CMP	ATT	PCT	PassYd	AvgYds/Pass	PassTD	INT	Rating	RushAtt	RushYds	AvgYds/Rush	QB Points Allowed
23.06	29.666667	44.666667	66.466667	322.333333	7.266667	2.000000	1.000000	93.253333	4.333333	17.333333	4.266667	14.066667

CORRELATION

	Fantasy Points
Fantasy Points	1.000000
CMP	0.418624
ATT	0.413010
PCT	0.427900
PassYd	0.440887
AvgYds/Pass	0.398332
PassTD	0.397824
INT	0.109928
Rating	0.452879
RushAtt	0.300473
RushYds	0.270769
AvgYds/Rush	0.273930
QB Points Allowed	0.084431



MODEL

- ▶ *Fantasy Points*

$$\begin{aligned} = & -0.0515(PCT) + 12.4197(INT) + 1.6585(YPR) + 0.0034(PCT)^2 \\ & + 0.0374(YPR)^2 - 2.997(INT)^2 - 0.0832(PCT)(INT) \\ & - 0.0137(PCT)(YPR) - 0.6841(INT)(YPR) \end{aligned}$$

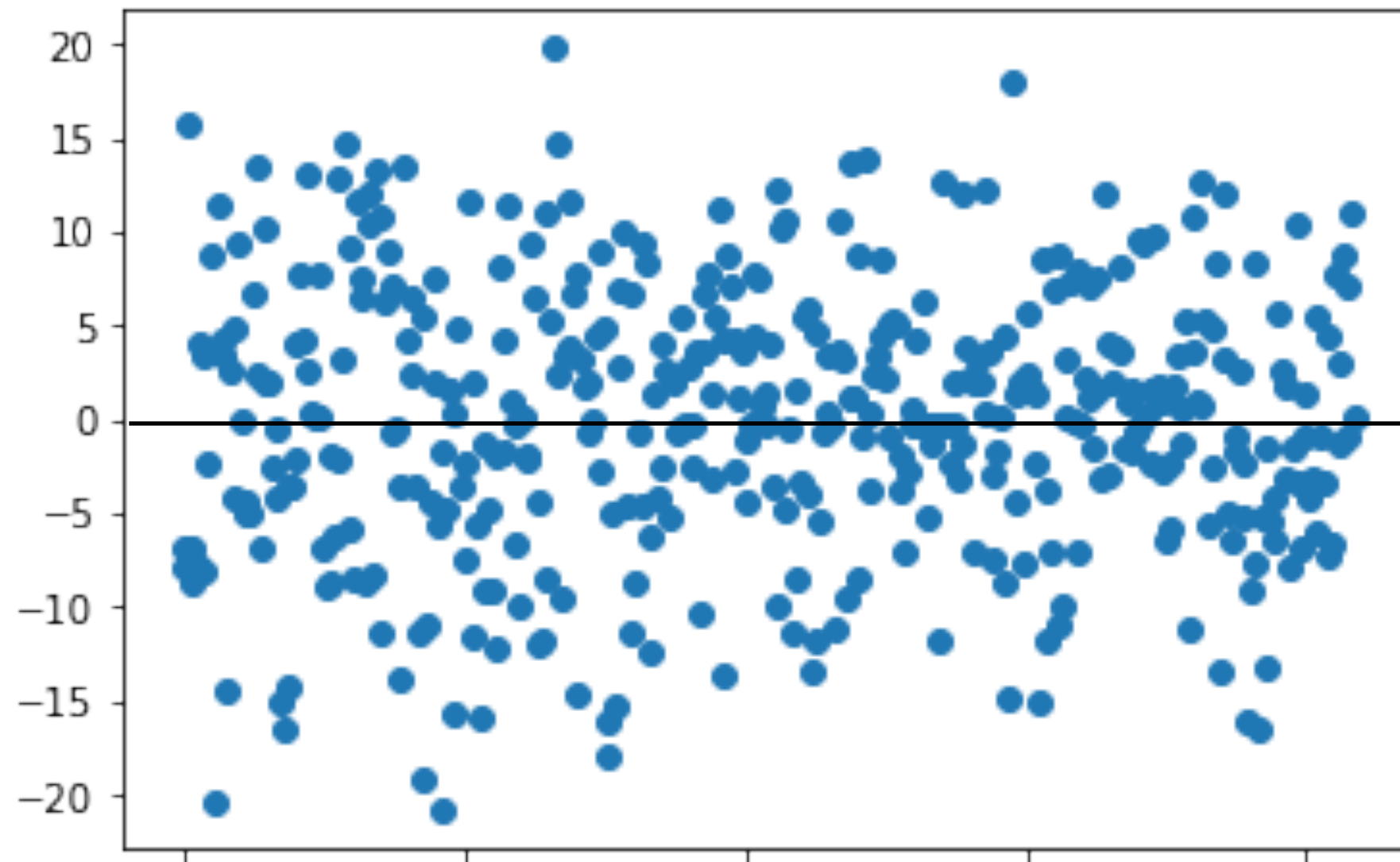
- ▶ R² metrics:

- ▶ Train: 0.231

- ▶ Test: 0.194

RESIDUALS

Residual Plot



Model d

REASONS FOR LOW R-SQUARED VALUES

- ▶ Touchdowns mean HIGH variance
- ▶ More outside factors involved

FUTURE WORK

- ▶ Other factors (weather, over/under, more stats, etc.)
- ▶ Other positions (RB, WR, TE)