LHL MIDTERM PROJECT

IMPACT OF PHYSICAL ATTRIBUTES ON 3-POINT SHOOTING IN THE NBA'S LATEST SEASON

Olatunbosun Ayetan and Richard Li

PROJECT FLOW

Data acquisition and preparation from recommend NBA stats API
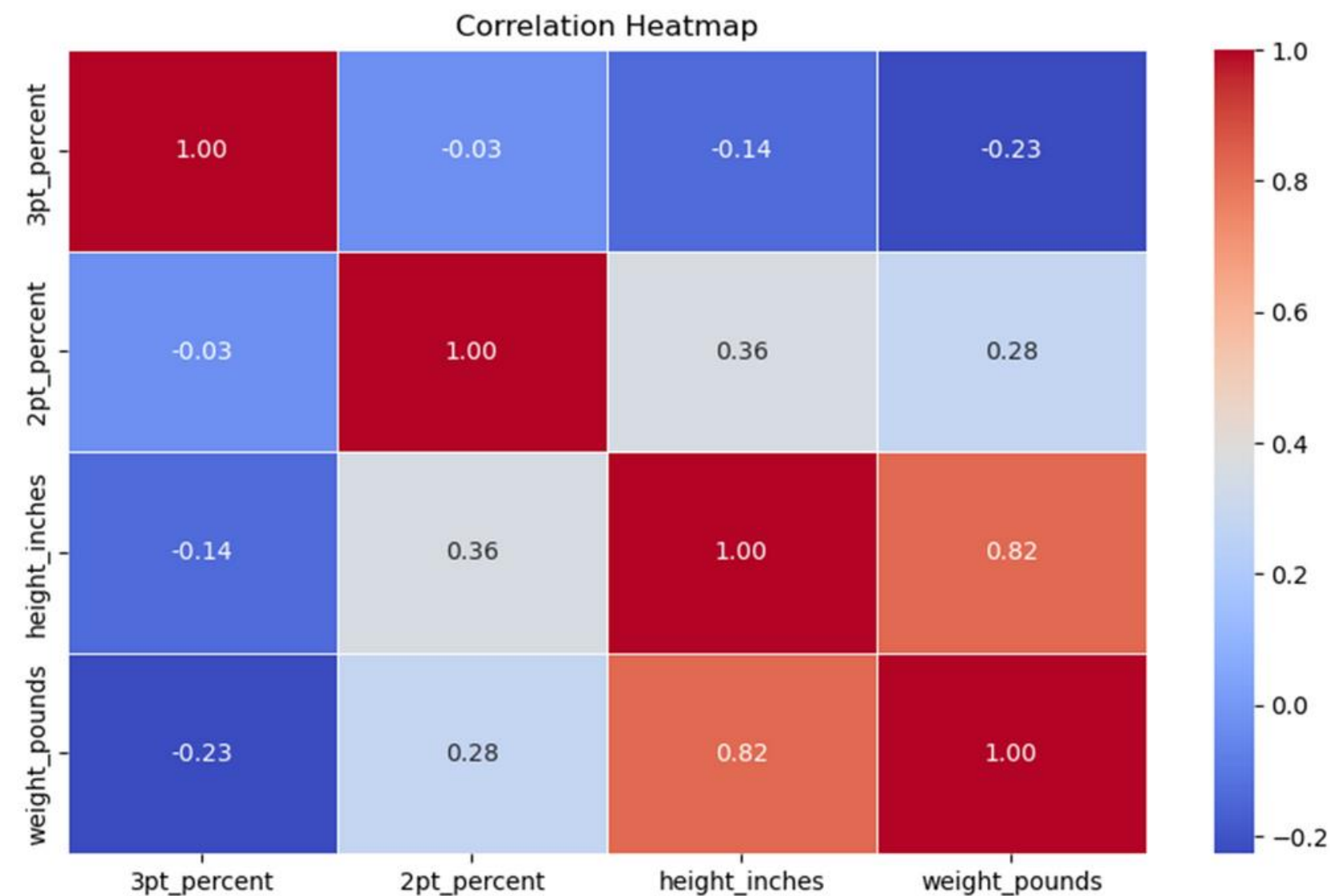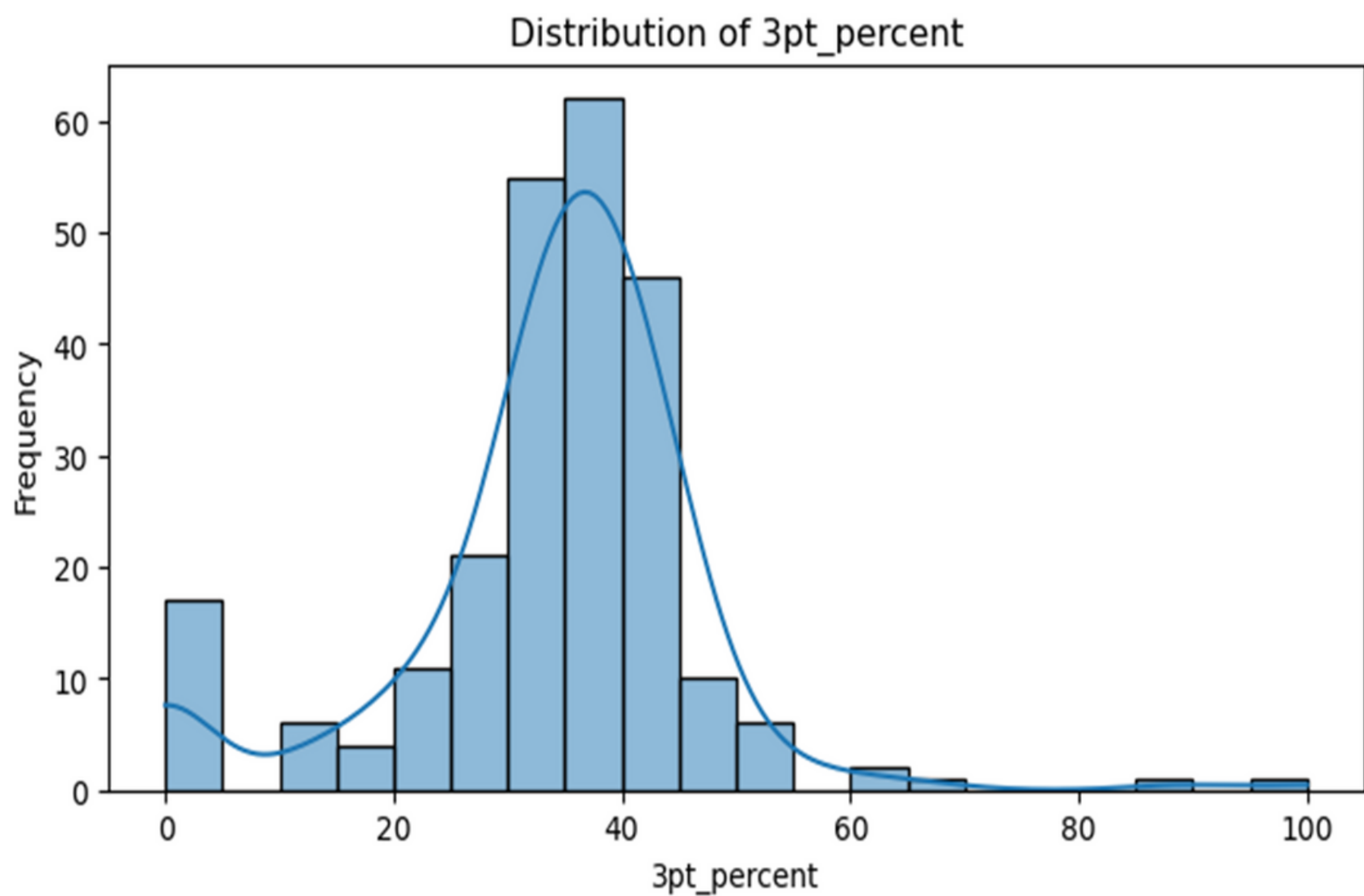
EDA on the resulting cleaned dataset

Created a multivariate linear regression model to study the impacts

Created a Tableau Dashboard to help visualize the potential conclusion

EXAMPLES EDA CHARTS

MODEL OUTPUT

```
Mean Squared Error (MSE): 206.69623497620728
R-squared (R2): 0.009676117030315456
                           OLS Regression Results
==============================================================================
Dep. Variable:             3pt_percent   R-squared:                       0.114
Model:                             OLS   Adj. R-squared:                  0.095
Method:                  Least Squares   F-statistic:                     6.080
Date:                 Wed, 06 Sep 2023   Prob (F-statistic):           0.000127
Time:                         14:07:24   Log-Likelihood:                -762.94
No. Observations:                  194   AIC:                             1536.
Df Residuals:                      189   BIC:                             1552.
Df Model:                            4
Covariance Type:             nonrobust
==============================================================================
                    coef     std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const           -19.6670      41.411     -0.475      0.635    -101.355      62.021
height_inches     1.0568       0.615      1.717      0.088      -0.157       2.271
weight_pounds    -0.1231       0.072     -1.709      0.089      -0.265       0.019
position_3.0     -3.4899       3.142     -1.111      0.268      -9.688       2.708
position_5.0    -14.4697       5.207     -2.779      0.006     -24.740      -4.199
==============================================================================
Omnibus:                      43.185   Durbin-Watson:                   2.113
Prob(Omnibus):                 0.000   Jarque-Bera (JB):              224.393
Skew:                          0.684   Prob(JB):                     1.88e-49
Kurtosis:                      8.088   Cond. No.                     1.08e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 1.08e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```
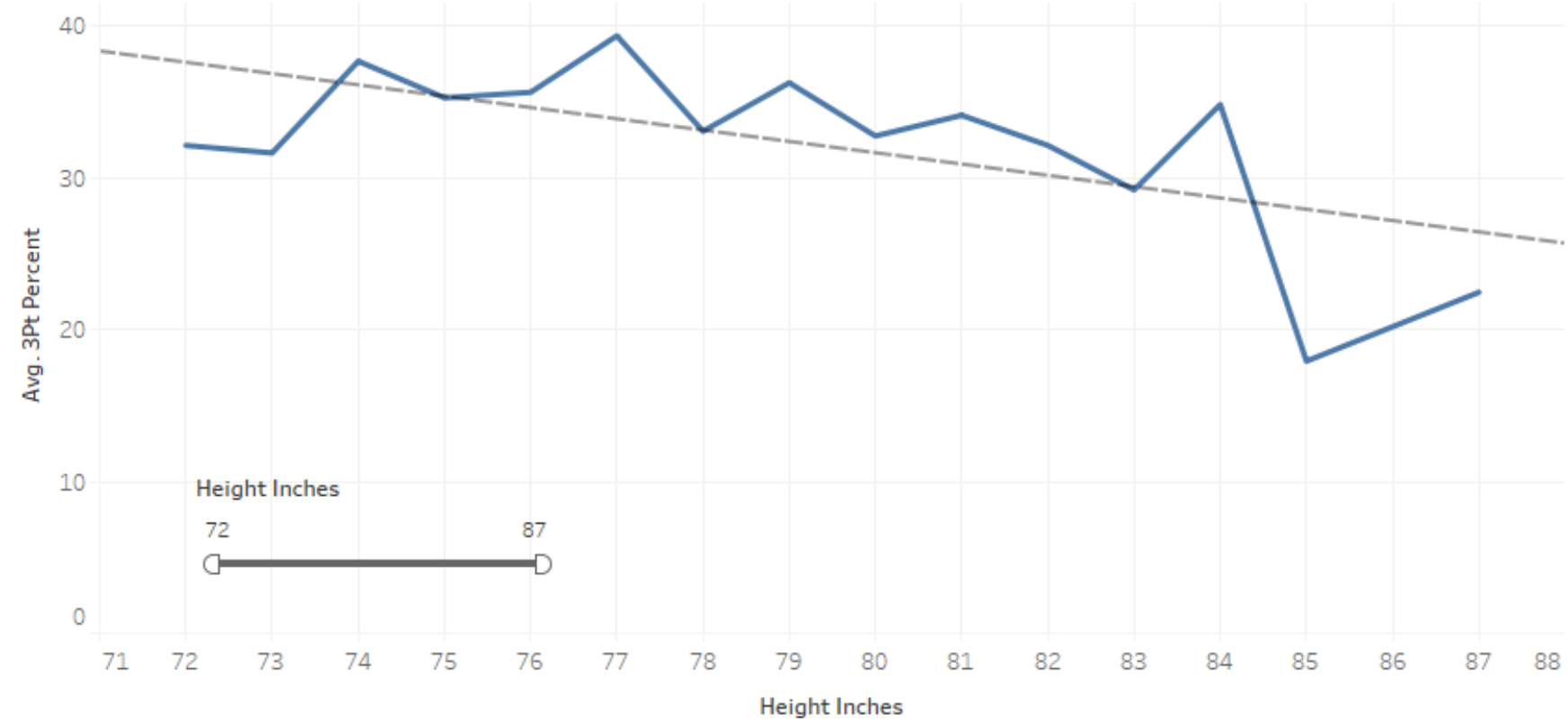
"*Significant p-value reveal the notable impact of player position, especially "Position 5," despite a weak linear relationship (low R-squared).*

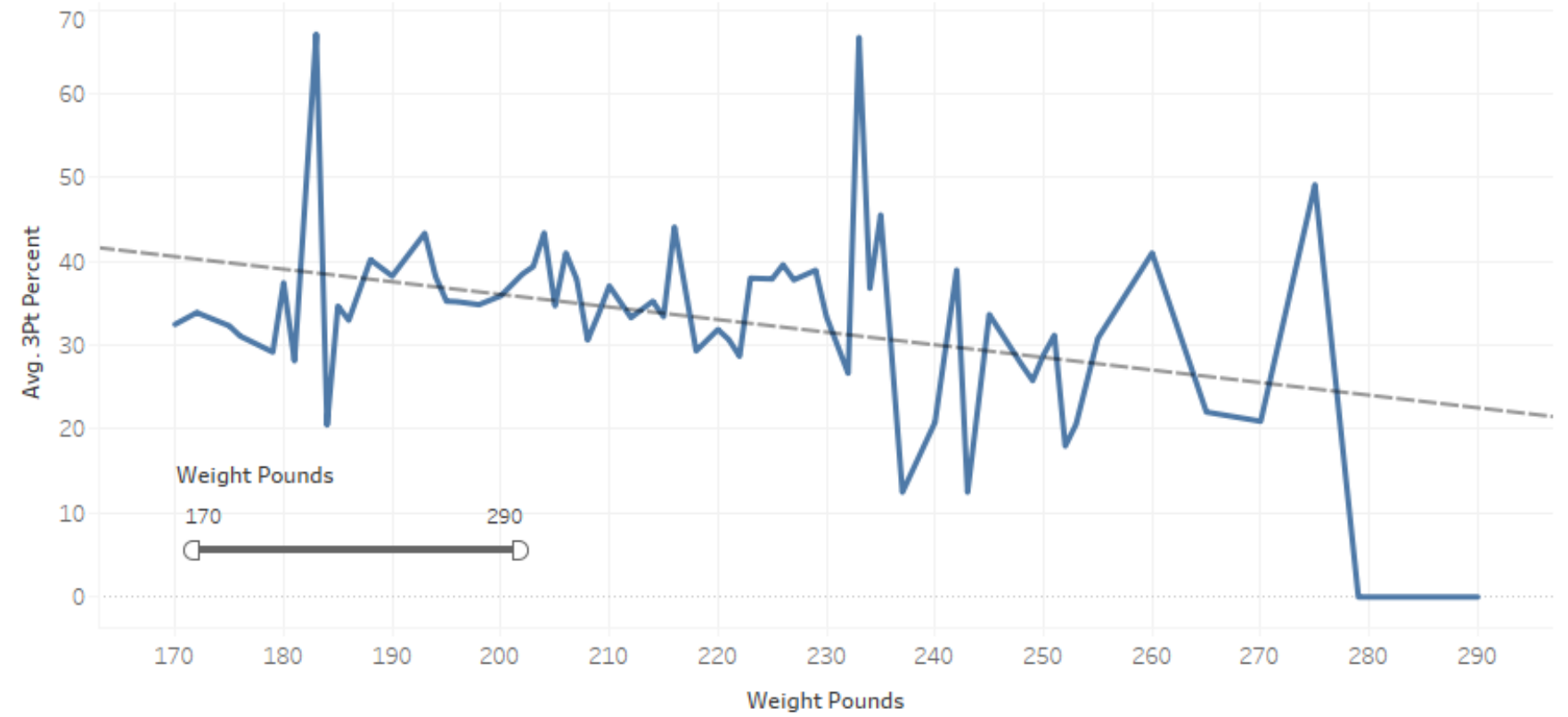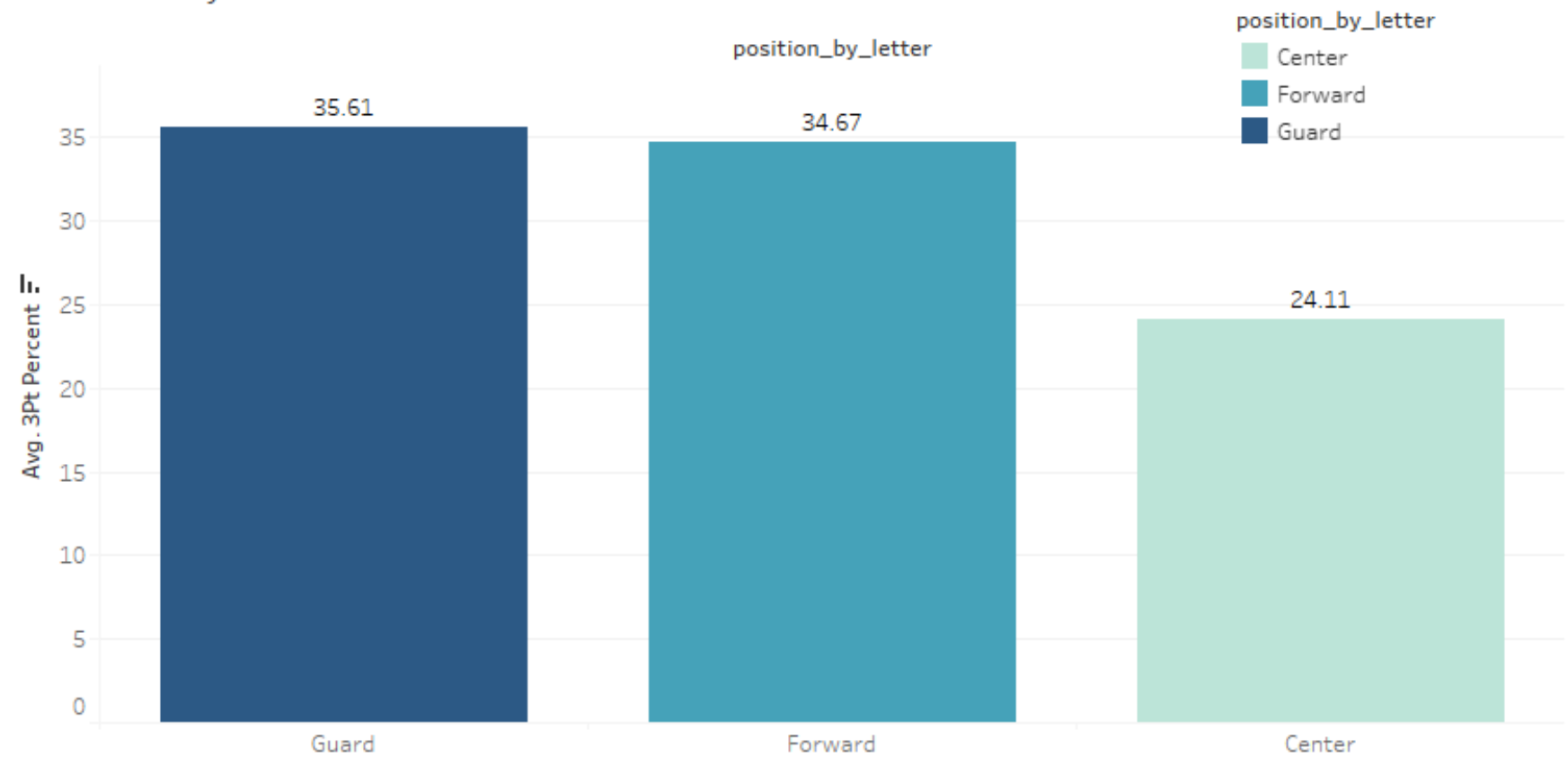*Position 5 has the greatest impact based on the coefficients*"
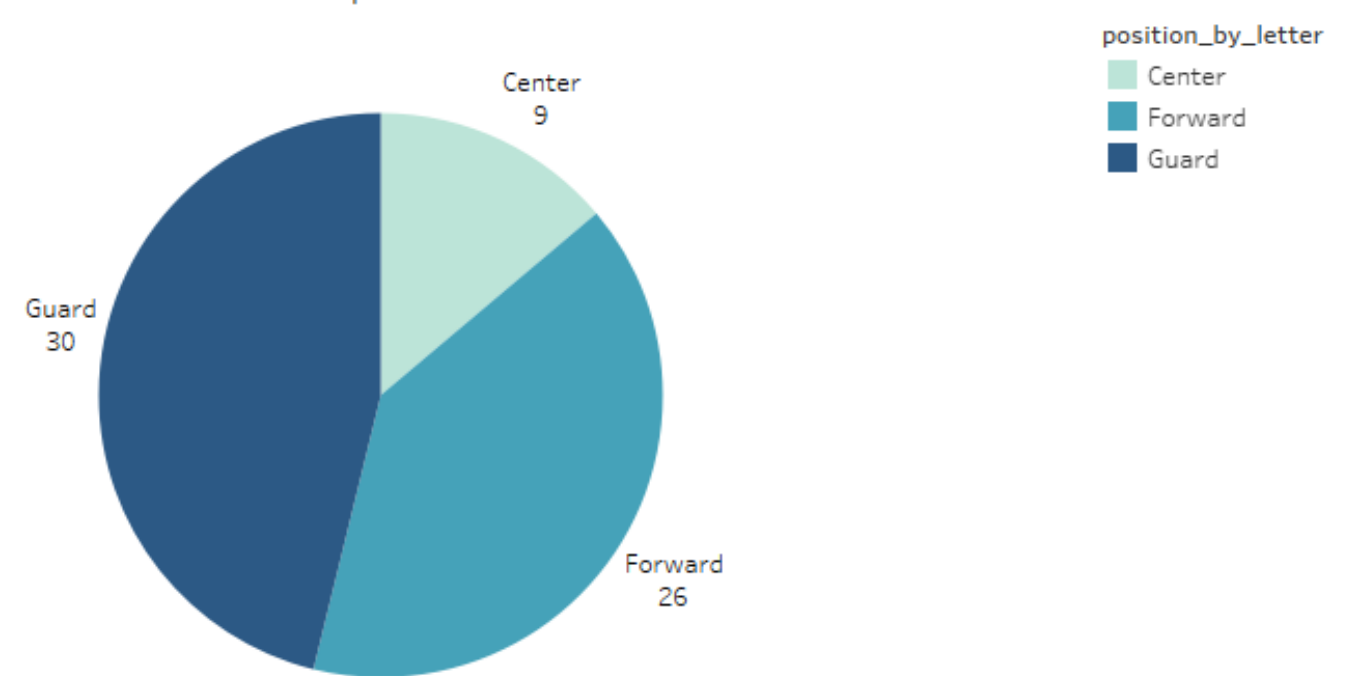
TABLEAU DASHBOARD

Height vs 3 Point %

Avg. 3Pt Percent vs Height Inches

Height Inches
72 — 87

Weight vs 3 Point %

Avg. 3Pt Percent vs Weight Pounds

Weight Pounds
170 — 290

3 Point % by Position

position_by_letter
- Center
- Forward
- Guard

Guard 35.61
Forward 34.67
Center 24.11

Number of Elite Shooters per Position

position_by_letter
- Center
- Forward
- Guard

Center 9
Guard 30
Forward 26

CHALLENGE

The biggest challenge was limited data in the API with many missing player entries, leading to a smaller dataset, possibly contributing to the weak regression model.

THANK YOU