

Project Luther

By Metis Data Science Associates LLC.

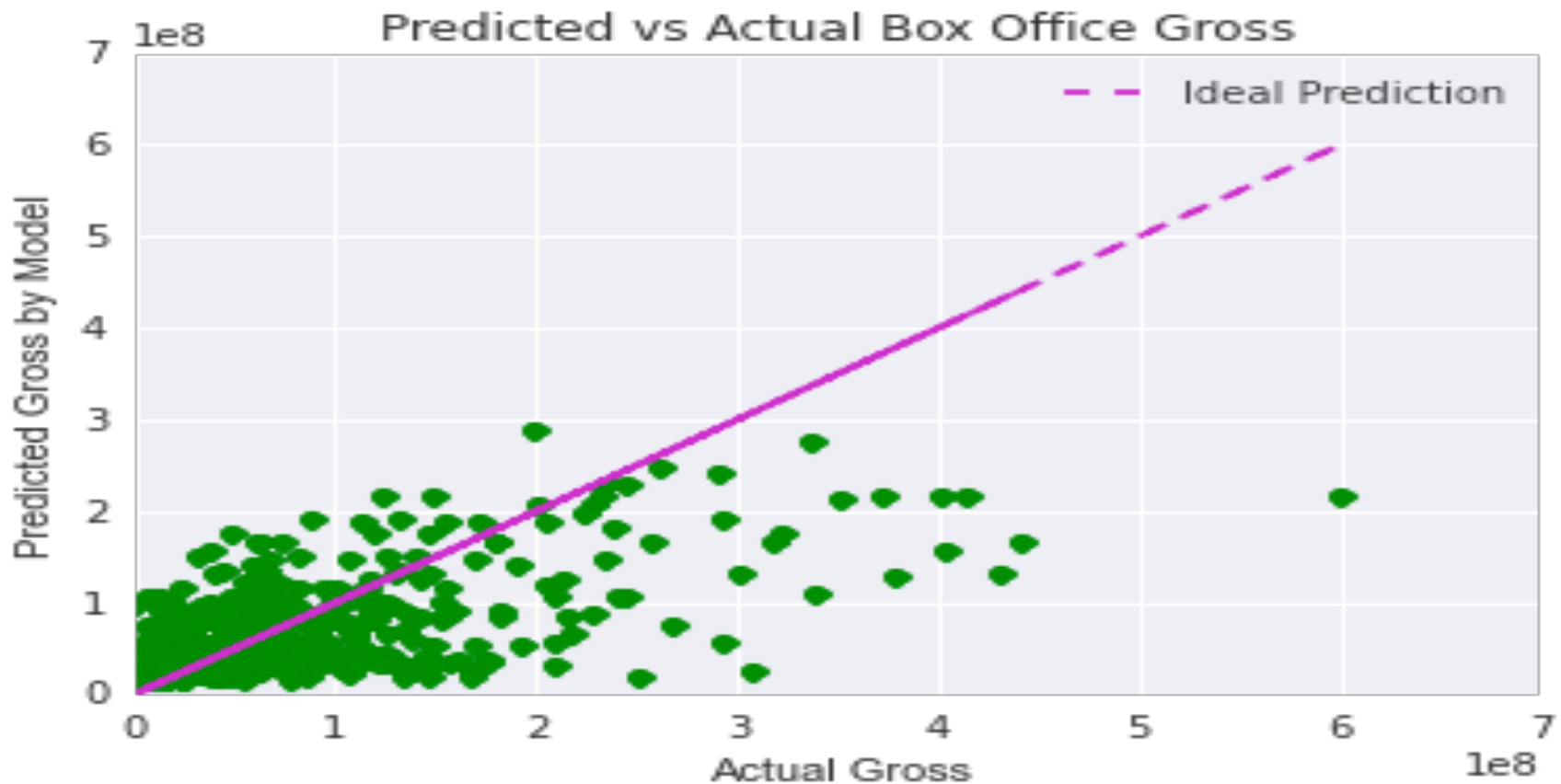
The Client's Request

- FundCo (the “Client”) is interested in investing in film projects and needs to model domestic gross revenue (“Box Office Gross”). They have suggested the use of historical data from boxofficemojo.com as the data source.

The **deliverable is a regression model** that allows them to predict domestic gross revenue.

Box Office Gross and Production budget

- **Production budget** is an intuitive model input.



The graph illustrates how the model fits the actual data.

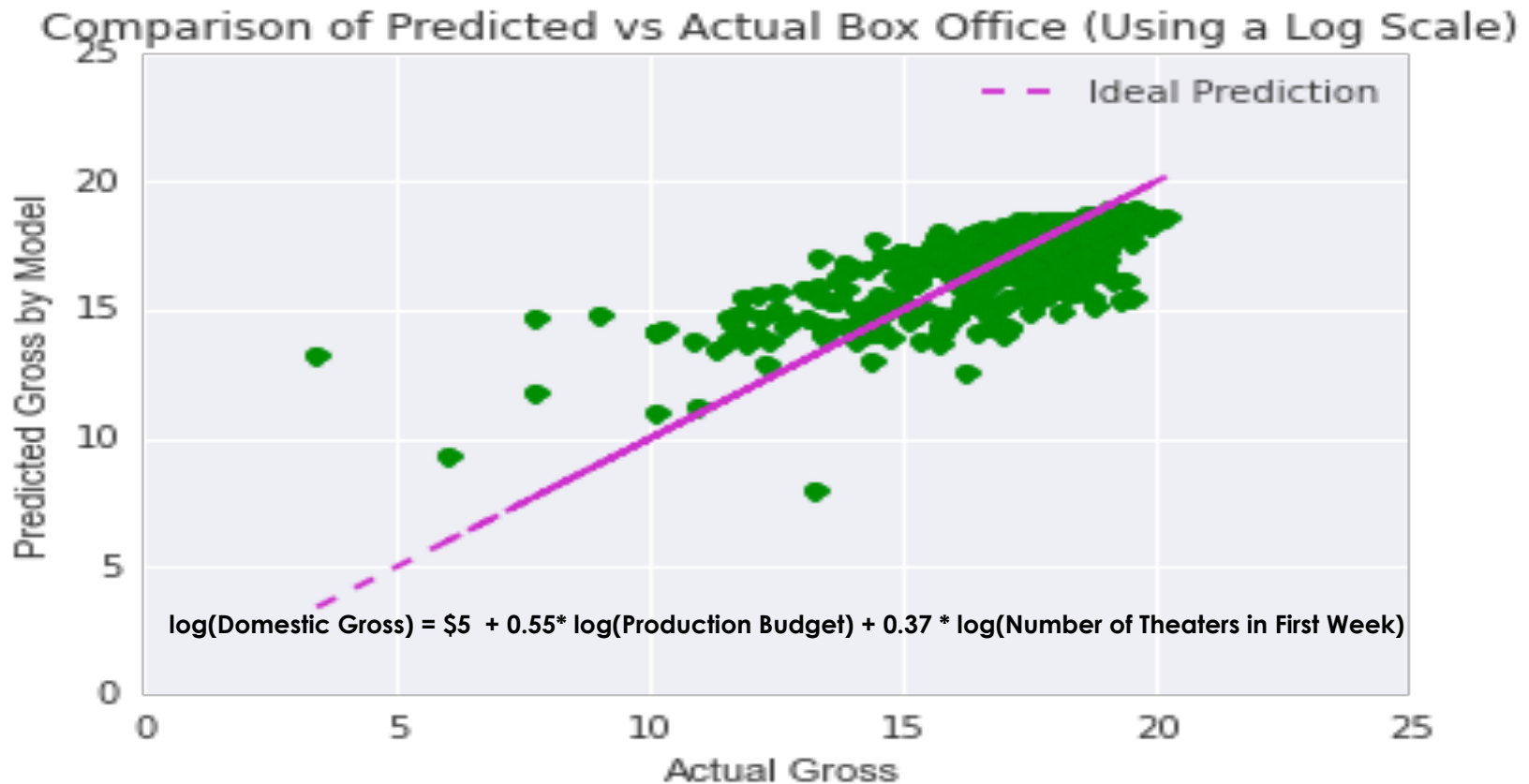
Box Office Gross with multiple input variables

- Using a single model input (**Production budget**) produces a regression model with an R^2 of 0.39.
- We infer that as the sole model input, **Production budget has limited predictive value.**
- The model can be improved by adding the **Number of theaters** that show the film in the first week.

NB: The R^2 coefficient of determination is a statistical measure of how well the regression line approximates the real data points. An R^2 of 1 indicates that the regression line perfectly fits the data.

The Model Solution

- A multiple regression model using both **Production budget** and **Number of theaters in first week** provides a better predictive framework



The model solution provides an improved fit data and higher. R^2 of 0.53.

The Model Solution

■ The Final Model

$$\log(\text{Domestic Gross}) = 5.0992 + 0.5498 * \log(\text{Production Budget}) \\ + 0.3727 * \log(\text{Number of Theaters in First Week})$$

The model solution provides an improved fit data and higher. R^2 of 0.53.

Thank you

Contact:

Metis Data Science Associates LLC.

79 Madison Ave

New York , NY

Tel: (212)-111-1111