New Hire Forecast

Modeling the CDM Team with Python

Multiple Linear Regression (MLR)

Simple Linear Regression

$$y = mx + b$$

Multiple Linear Regression

$$y = m_1x_1 + m_2x_2 + \cdots m_nx_n + b$$

Where:

x_1

x_2

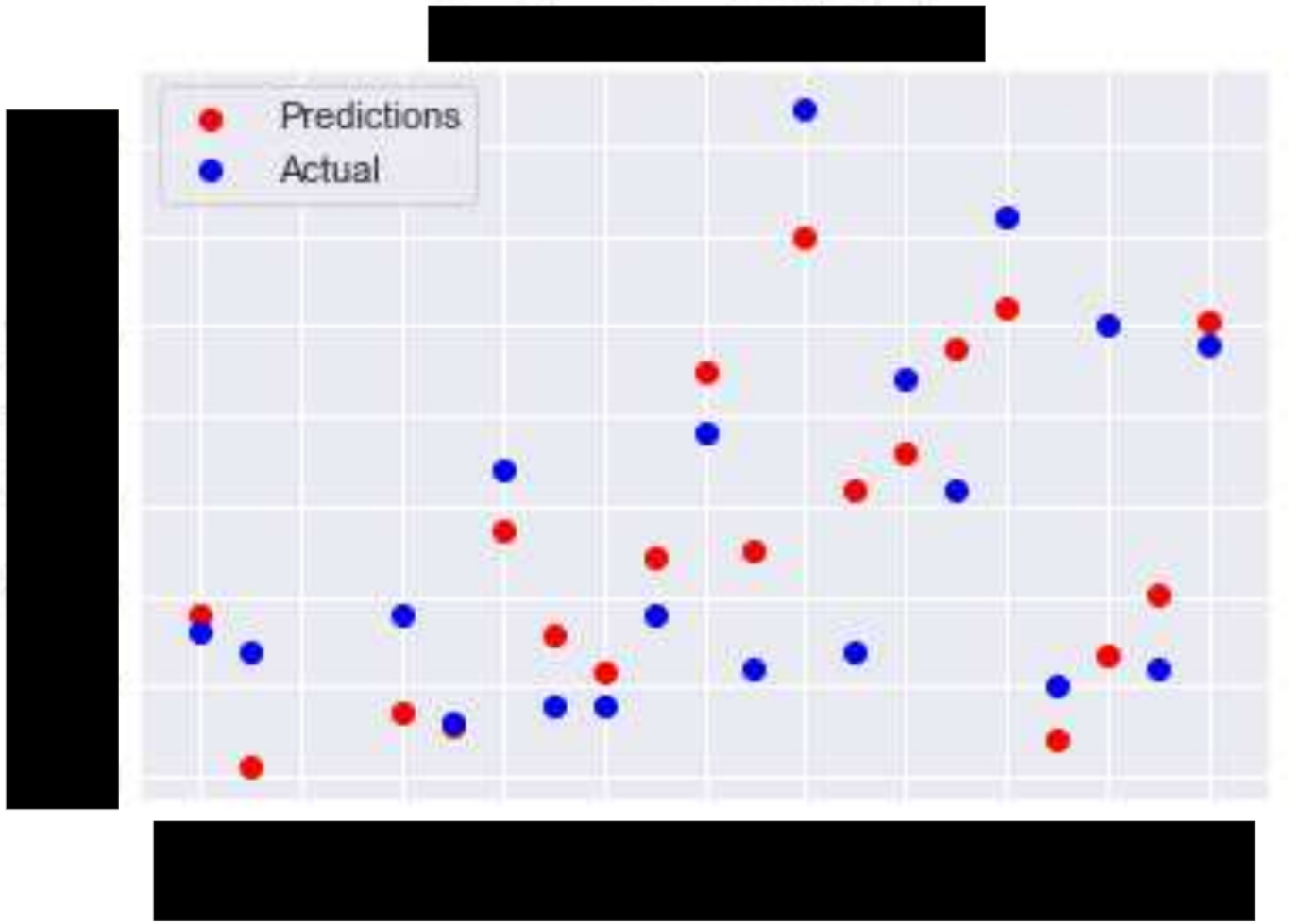
y

m_n

b

https://jakevdp.github.io/PythonDataScienceHandbook/05.06-linear-regression.html

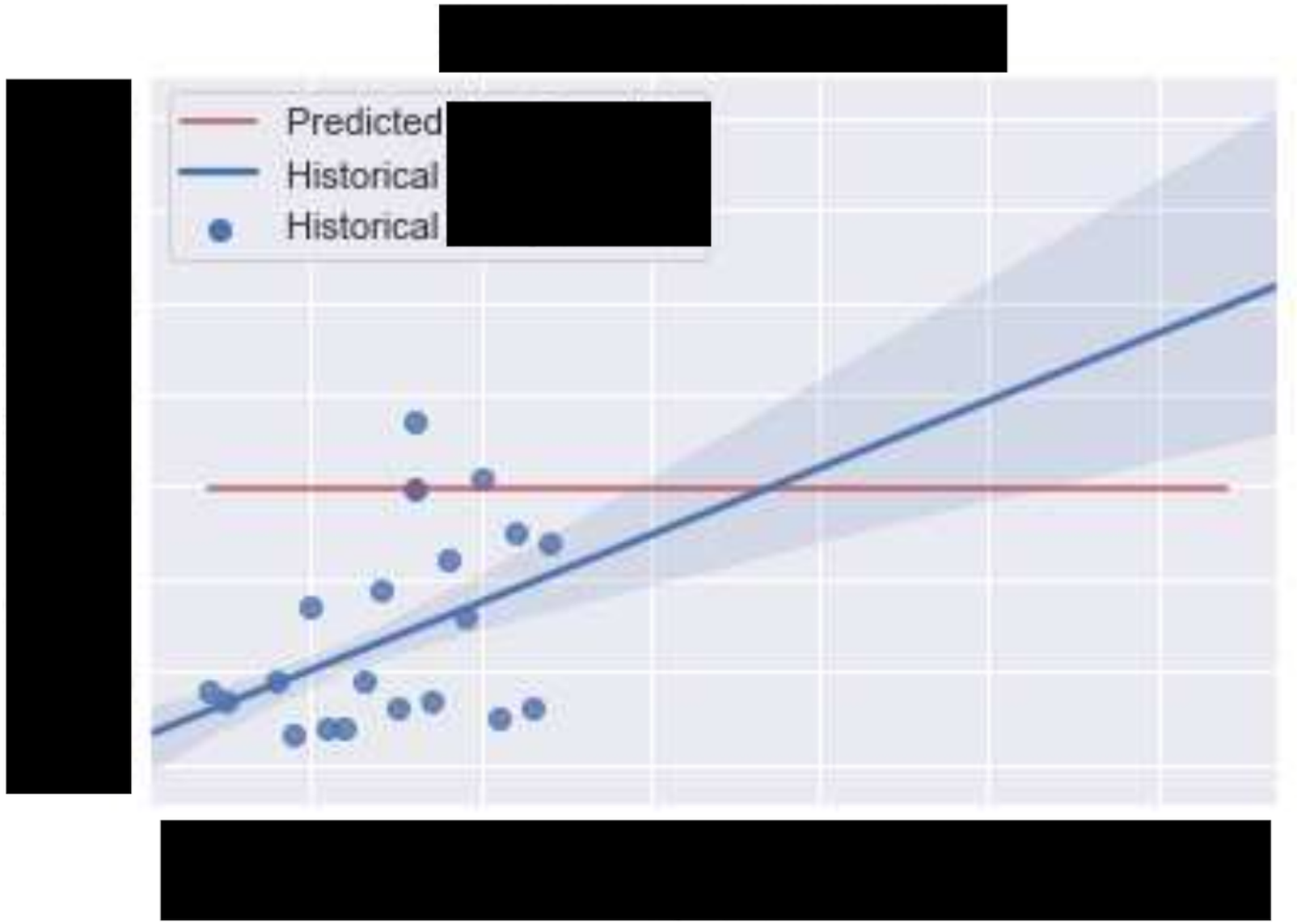
MLR Model



- Each red dot is a calculated $\hat{y}$
- Point approximations vary up to +/- ■
- Aside from individual points, the model variance also trends with the data
  - Can calculate and graph this – but that's not our goal here
- Many purposes
  - In this context, we use it for a theoretical weighted maximum

Projecting into the future: MLR and LR

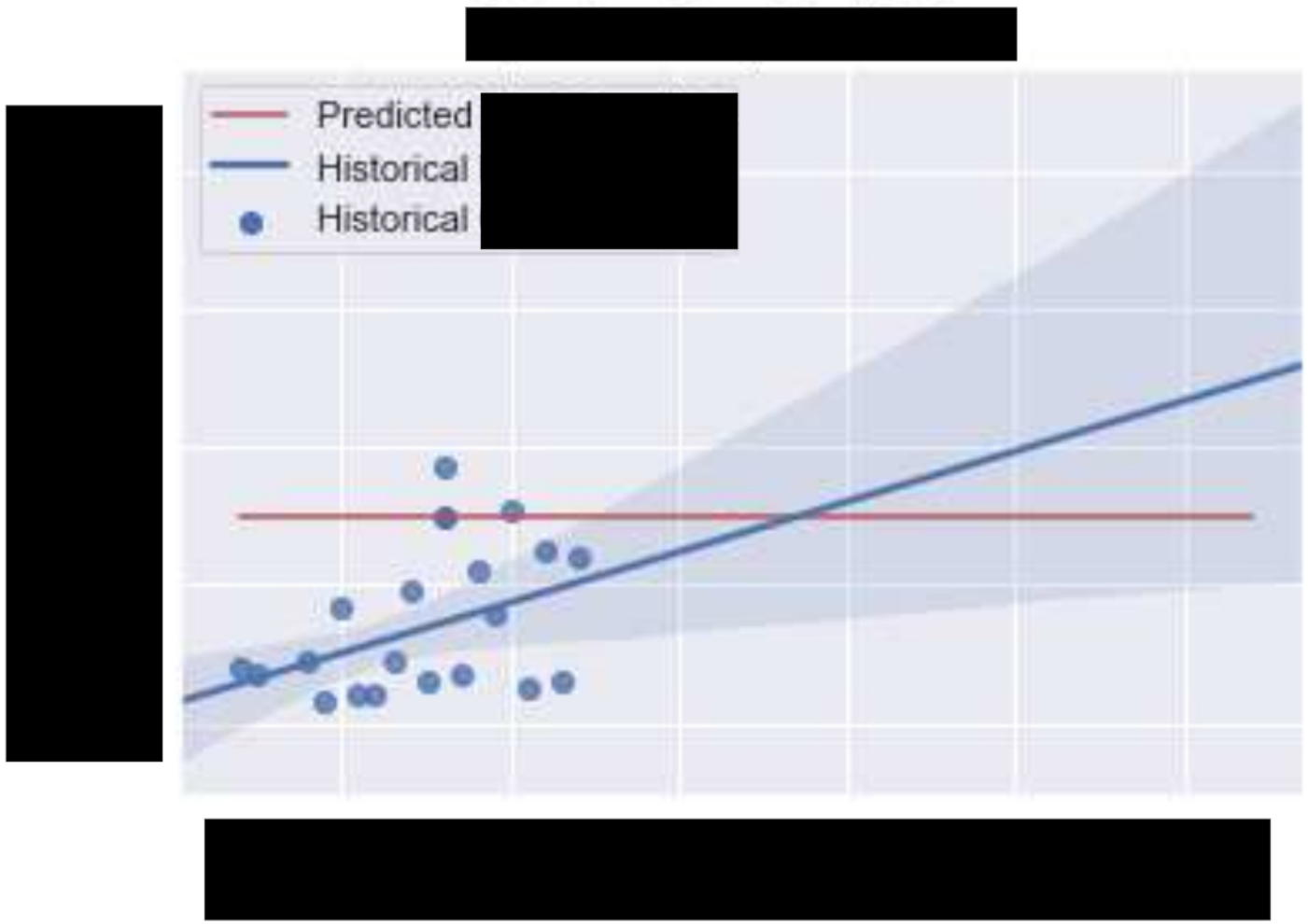
68% confidence interval (σ=1)



- Key Data
  - Intersection at [redacted]
  - 68% Upper Bound at [redacted]
  - 68% Lower bound at [redacted]
- "Bootstrapped" Confidence Interval
  - Non-analytical method of computing the confidence interval https://en.wikipedia.org/wiki/Bootstrapping_(statistics)
  - Embedded within the Seaborn library https://seaborn.pydata.org/generated/seaborn.regplot.html

Projecting into the future: MLR and LR

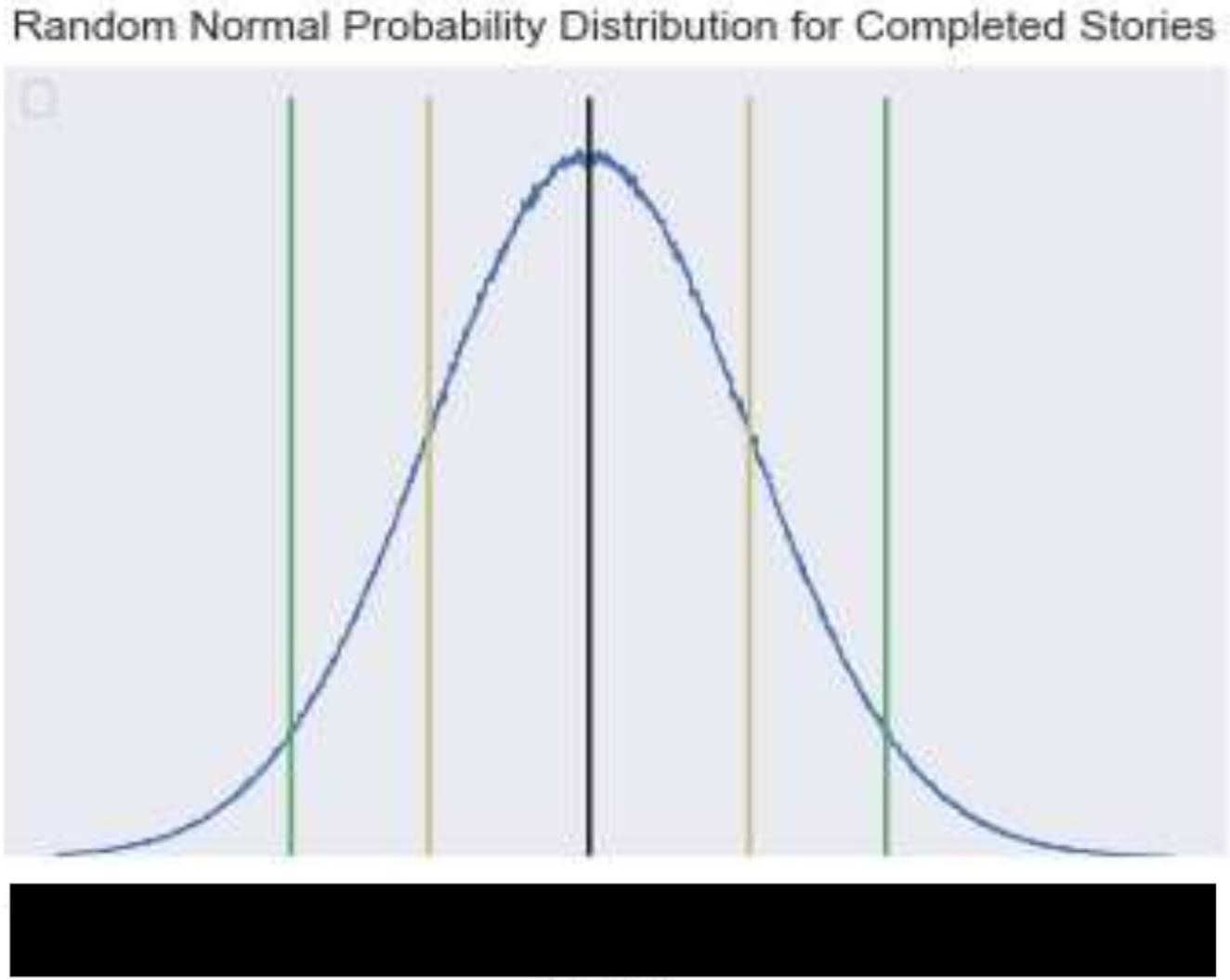
95% confidence interval (σ=2)



- Key Data
  - Intersection at
  - 95% Upper Bound at
  - 95% Lower bound at

Another view:

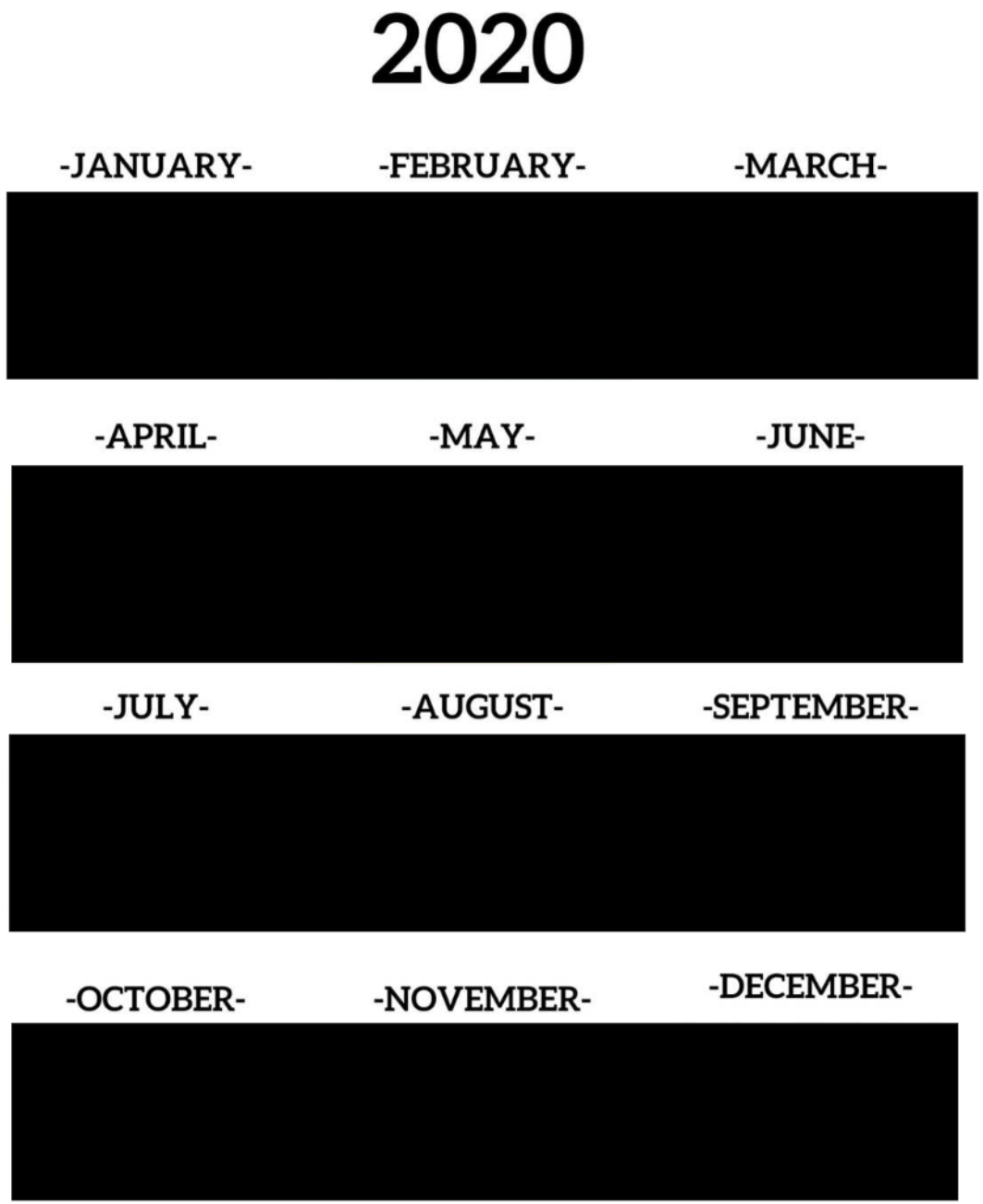
Random Normal Probability Distribution @ [redacted]



- Key Data
  - Center (mean) at [redacted]
  - 68% Upper Bound at [redacted]
  - 68% Lower bound at [redacted]
  - 95% Upper Bound at [redacted]
  - 95% Lower bound at [redacted]
- Central Limit Theorem (CLT)
  - As explained by the CLT, the bootstrapped confidence interval is *approximately* symmetric around the point of interest; thus it is not out-of-the-question to model it as a representation of the actual mean and standard deviation https://en.wikipedia.org/wiki/Central_limit_theorem

Calendar View

- Key Data
  - Intersection: [redacted]
  - 68% upper boundary: [redacted]
  - 95% lower boundary: [redacted]
  - 95% upper boundary: [redacted]



Disclaimers

- [redacted]
  - [redacted]
  - [redacted]
    - [redacted]
- The real world is more complicated
  - Business changes [redacted]
  - Worker efficiency
  - Actual maximum

Next steps

- [redacted]
  - [redacted]
- [redacted]
- [redacted]

Thank you for listening ☺