# REGRESSION ANALYSIS OF CAR AUCTION BIDS

MANVEER SADHAL

OCT 1, 2021

# QUESTION

What is the most you should bid to win an online car auction without overpaying?

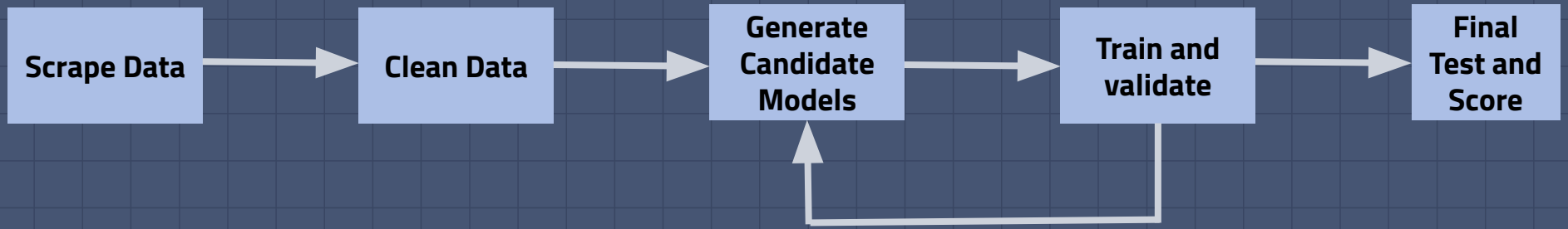
# DATA AND SCOPE

- Data
  - BringATrailer.com completed auctions
- Scope
  - BMW M3 model years 1994-2013
  - 1,121 auctions evaluated from Dec 2016 to present
  - 908 records remained after cleaning

# TOOLS

- Web Scraping
  - Selenium
  - BeautifulSoup
- Data Cleaning
  - Pandas
- Modeling
  - SciKit-Learn
- Visualization
  - Matplotlib
  - Seaborn

# PROCESS

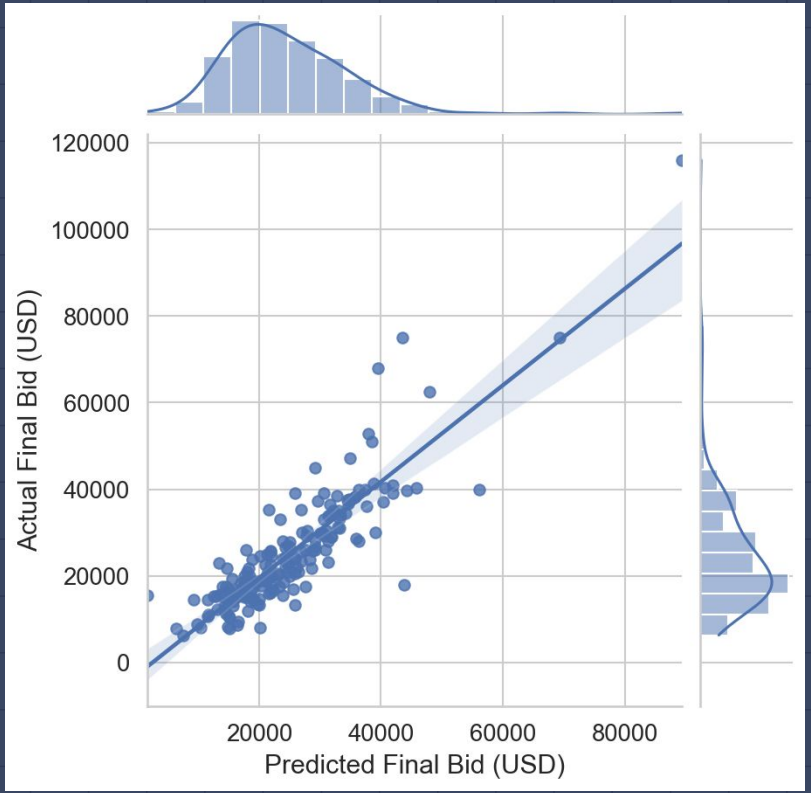



Add/remove/transform model features

Try different models

# RESULTS

- Linear Model Metrics
  - $R^2$: 0.76
  - Mean Absolute Error: $4,377
  - Root Mean Squared Error: $6,585



# RESULTS

Most Influential

Least Influential

| Feature | Standardized Coefficient |
|---|---|
| log(Mileage) | -6781 |
| Auction Close Date | 3028 |
| Model Year | 2844 |
| Not a Limited Edition | -2499 |
| Body Style | 2421 |
| Auction Reserve Met | 2178 |
| Manual Transmission | 1991 |

# LIMITATIONS

- Model is less accurate at extremes (cars with very high and very low final bids). Not enough data points in these regions to capture the trend.
- Sharp rise in used car prices in 2020-2021. Trend may not persist in upcoming years.

# CONCLUSIONS

- Mileage is the overwhelming driver of used car price
- The value of used cars has increased over recent years
- Newer cars are generally more expensive
- Special editions of the BMW M3 (e.g. lightweight) hold their value exceptionally well
- Convertible tops and automatic transmissions are less desirable in the M3

# FURTHER WORK

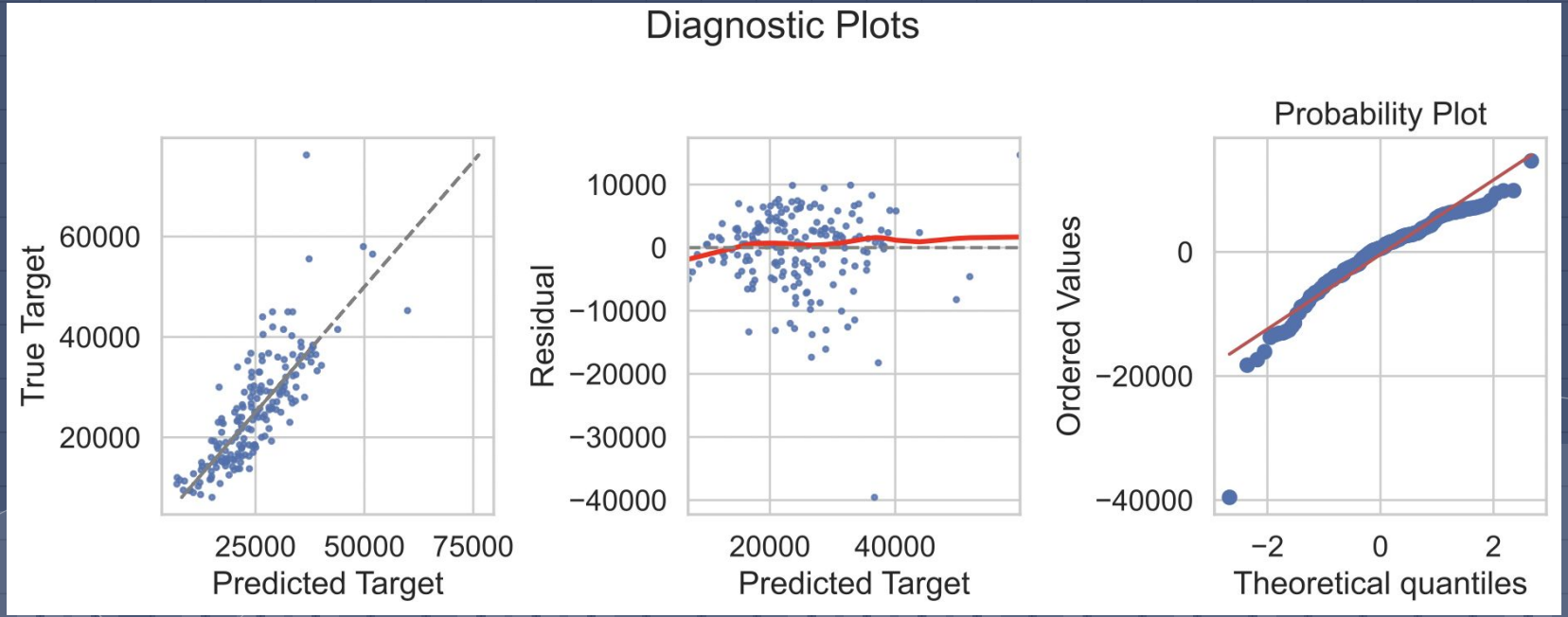
- Collect additional data
  - Other auction websites
  - Individual vehicle history (e.g. accidents, number of owners)
- Use alternate models:
  - Time series to predict trends
  - Tree model to adapt to nonlinear relationships

# THANKS!

**Any questions?**

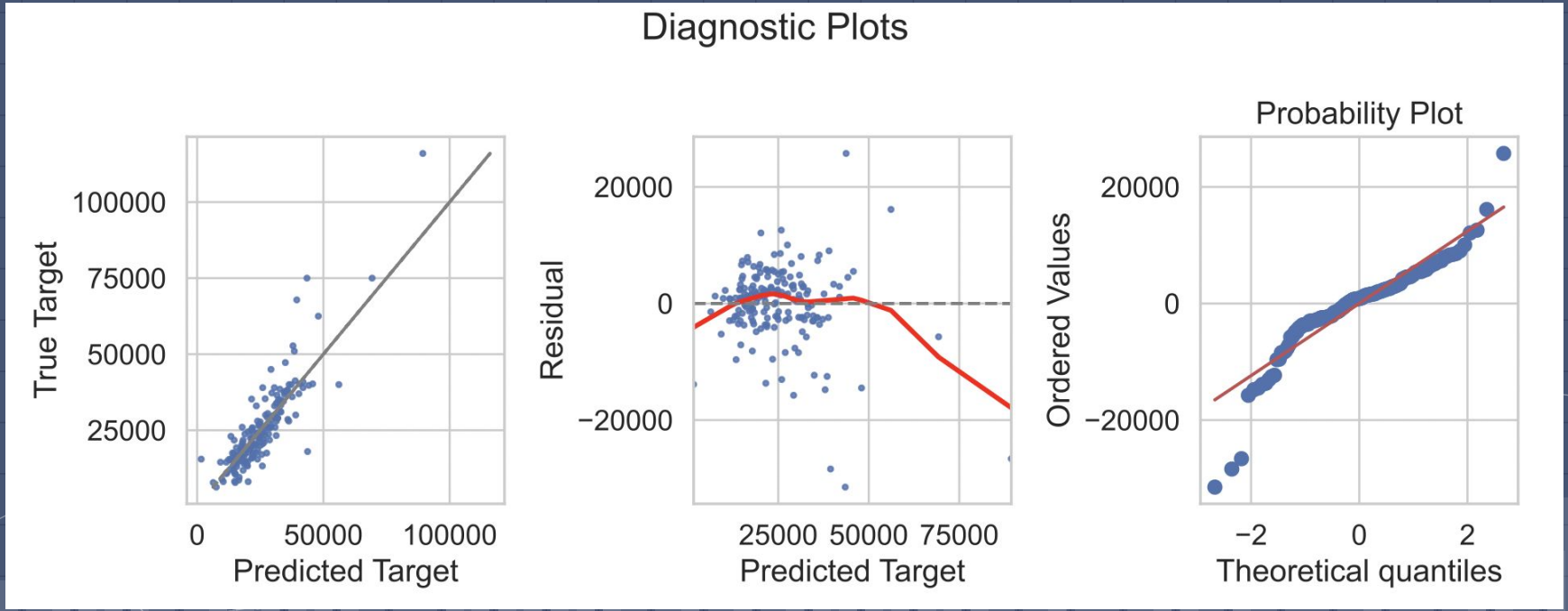
# Appendix

- Final validation



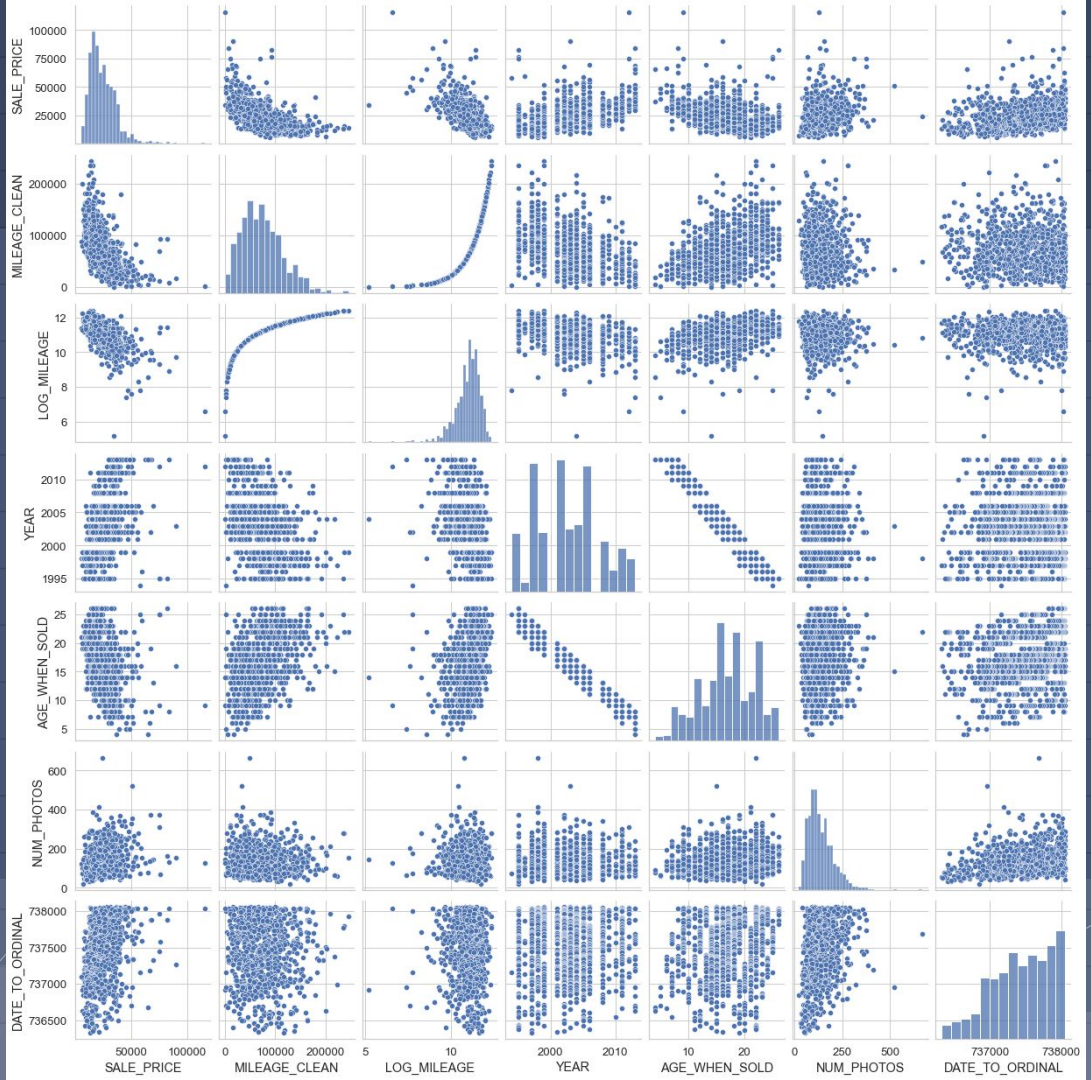
# Appendix

- Test



# Appendix

| | SALE_PRICE | MILEAGE_CLEAN | LOG_MILEAGE | YEAR | AGE_WHEN_SOLD | NUM_PHOTOS | DATE_TO_ORDINAL |
|---|---|---|---|---|---|---|---|
| **SALE_PRICE** | 1.000000 | -0.583189 | -0.615759 | 0.472826 | -0.405725 | 0.218655 | 0.278991 |
| **MILEAGE_CLEAN** | -0.583189 | 1.000000 | 0.833499 | -0.417688 | 0.412189 | -0.031762 | -0.027489 |
| **LOG_MILEAGE** | -0.615759 | 0.833499 | 1.000000 | -0.383544 | 0.385722 | -0.031190 | 0.002676 |
| **YEAR** | 0.472826 | -0.417688 | -0.383544 | 1.000000 | -0.969747 | -0.044772 | 0.089513 |
| **AGE_WHEN_SOLD** | -0.405725 | 0.412189 | 0.385722 | -0.969747 | 1.000000 | 0.118754 | 0.149821 |
| **NUM_PHOTOS** | 0.218655 | -0.031762 | -0.031190 | -0.044772 | 0.118754 | 1.000000 | 0.326935 |
| **DATE_TO_ORDINAL** | 0.278991 | -0.027489 | 0.002676 | 0.089513 | 0.149821 | 0.326935 | 1.000000 |


SALE_PRICE
MILEAGE_CLEAN
LOG_MILEAGE
YEAR
AGE_WHEN_SOLD
NUM_PHOTOS
DATE_TO_ORDINAL

# CREDITS

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by SlidesCarnival
- Photographs by Unsplash