0.0.1 Question 2c: Verify Outcome

Did the candidate win or lose the election? Verify with election outcome.

ANSWER

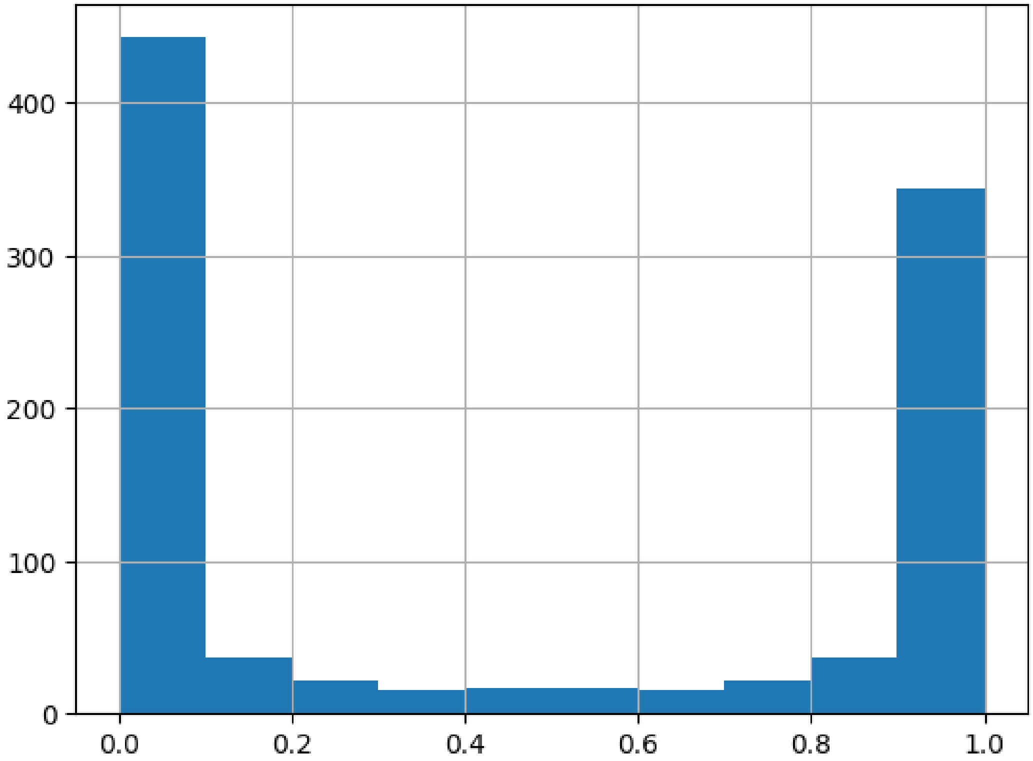
Kevin Yoder, the rising candidate won the election. Sharice Davids hte falling candidate did not win the election. This alligns with the results we had from earlier.

0.0.2 Question 3a: Prediction Histogram

Make a histogram showing the predicted win probabilities *on the morning of the election.* Again, restrict yourself to only the `classic` predictions.

```
In [14]: election_data[(election_data['forecast_type'] == 'classic') & (election_data['forecast_date']
```

```
Out[14]: <Axes: >
```



0.0.3 Question 3b: Prediction difficulty

Are most house elections easy to forecast or hard to forecast? State your reasoning.

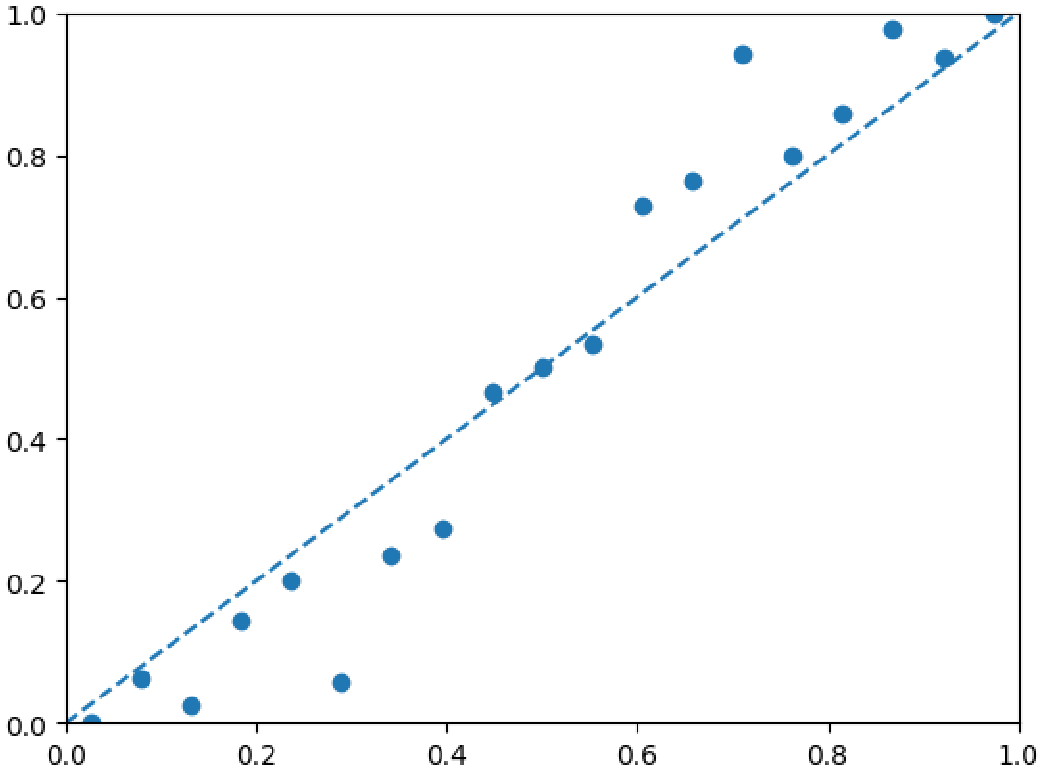
ANSWER

Based on the histogranm the house predicitons are easy to predict. The highest predicted win probabilities are closer to 0 or 1 which means the model

0.0.4 Question 4c: Visualize Results

Now make a scatterplot using `midpoints` as the x variable and `fraction_outcome` as the y variable. Draw a dashed line from `[0,0]` to `[1,1]` to mark the line y=x.

```
In [24]: import matplotlib.pyplot as plt
         x = midpoints
         y = fraction_outcome
         plt.scatter(x, y)
         plt.plot([0, 1], [0, 1], '--')
         plt.axis(xmin = 0, xmax = 1, ymin = 0, ymax = 1)
         plt.show()
```

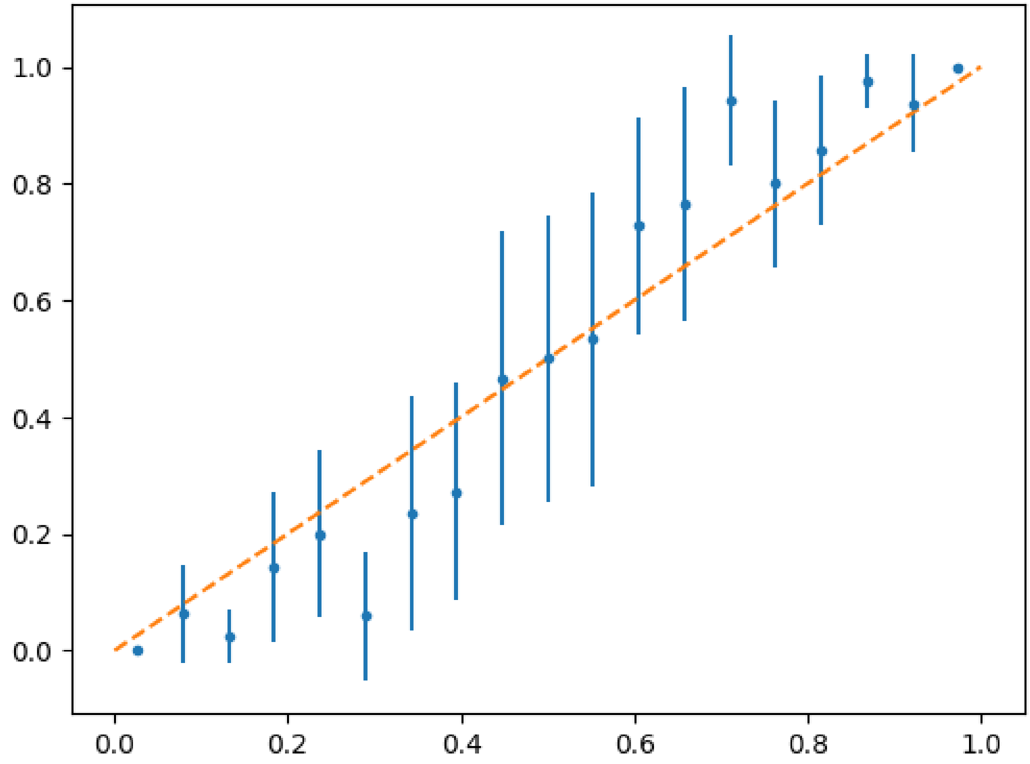


0.0.5 Question 5b: Visualize Error Bars 1

Use `plt.errorbar` to create a new plot with error bars associated with the actual fraction of wins in each bin. Again add a dashed y=x line. Set the argument `fmt='.'` to create a scatterplot with errorbars.

```
In [27]: # Plotting code below
         plt.errorbar(midpoints, election_agg['mean'].values, yerr=election_agg['err'].values, fmt='.')
         plt.plot([0, 1], [0, 1], '--')
```

```
Out[27]: [<matplotlib.lines.Line2D at 0x7f30bdf0b340>]
```



0.0.6 Question 5d: Understanding Confidence Intervals

Are the 95% confidence intervals generally larger or smaller for more confident predictions (e.g. the predictions closer to 0 or 1). What are the factors that determine the length of the confidence intervals?

ANSWER

The confidence intervall is small than for the predicitions. This is because the variance is inderectly proportional to the varaiance, so the smaller sample size will make the confidence larger.

0.0.7 (PSTAT 234) Question 5f. Visualize Error Bars 2

By now, we have a distribution of success probabilities saved in `bootstrap_election_agg`. We can compute empirical error bars from 2.5% and 97.5% quantiles. Write function named `bootstrap_error_bars` that can be used to calculate the following columns:

- `mean`: mean of probabilities of success
- `err_low`: low point of the error bars
- `err_high`: high point of the error bars

Funtion `bootstrap_error_bars` is to be called by using `bootstrap_election_100_agg.apply(bootstrap_error_bars, ...)`.

```
In [ ]: def bootstrap_error_bars(x):
            out = pd.Series([x.mean(), x.mean()-x.quantile(0.025), x.quantile(0.975)-x.mean()],
                             index=['mean', 'err_low', 'err_high'])
            return(out)
```

0.0.8 (PSTAT 234) Question 5g: Interpreting the Results

Are the 95% confidence intervals generally larger or smaller for more confident predictions (e.g. the predictions closer to 0 or 1). What are the factors that determine the length of the error bars?

Compare and contrast model-based error bars and empirically obtained error bars. What are the advantages and disadvantages of these two approaches?

Type your answer here, replacing this text.