PROJECT NO. 4

01

02

TEAM 2

UNVEILING NFL INSIGHTS THROUGH STATISTICS, SPREADS AND DATA ANALYTICS

GAME CHANGER

TEAMWORK

THE ROSTER

LOUIS FERRANTE

DATA CLEANING

JONAH FOEDAY

VISUAL-IZATIONS

BRITTNEY HILL

VISUAL-IZATIONS

TEAMWORK

THE ROSTER

CONTINUED

ALYSA SCHOENFELDER

VISUALIZATIONS

PRESENTATION

RYLAND WHELISS

MACHINE LEARNING

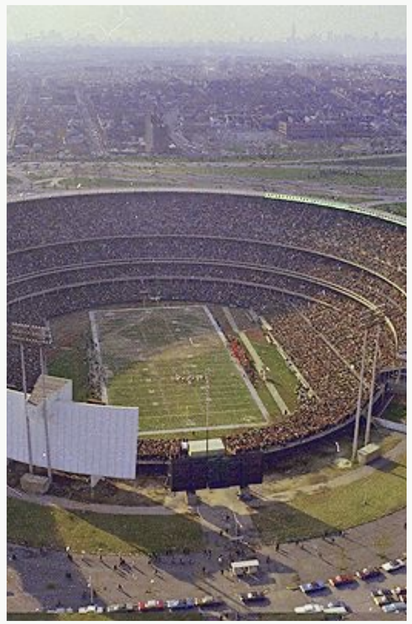
GAME PLAN

DEVELOP A MACHINE LEARNING MODEL TO FORECAST THE WINNING NFL FOOTBALL TEAM

DATA DRAFT

GAME / SPREAD

01



STADIUMS

02

TEAMS

03

CONDITIONING

```
def new_col(row):
    if row['score_home'] == row['score_away']:
        val = 2
    elif row['score_home'] > row['score_away']:
        val = 1
    else:
        val = 0
    return val
```

1

```
df_big_data['winner(0_away,1_home,2_tie)'] = df_big_data.apply(new_col, axis=1)

df_big_data.head()
```

```
bigoledata = df_big_data.merge(df_teams, left_on = 'team_home', right_on = 'team_name', how = 'left', suffixes = (None, '_h
bigoledata.head()
```

	schedule_date	schedule_season	schedule_week	schedule_playoff	team_home	score_home	score_away	team_away	team_favorite_id
0	9/2/1966	1966	1	False	Miami Dolphins	14.0	23.0	Oakland Raiders	NaN
1	9/3/1966	1966	1	False	Houston Oilers	45.0	7.0	Denver Broncos	NaN
2	9/4/1966	1966	1	False	San Diego Chargers	27.0	7.0	Buffalo Bills	NaN
3	9/9/1966	1966	2	False	Miami Dolphins	14.0	19.0	New York Jets	NaN
4	9/10/1966	1966	1	False	Green Bay Packers	24.0	3.0	Baltimore Colts	NaN

5 rows × 26 columns

2

```
humungodata = bigoledata.merge(df_teams, left_on = 'team_away', right_on = 'team_name', how = 'left', suffixes = ('_home',
humungodata.head()
```

3

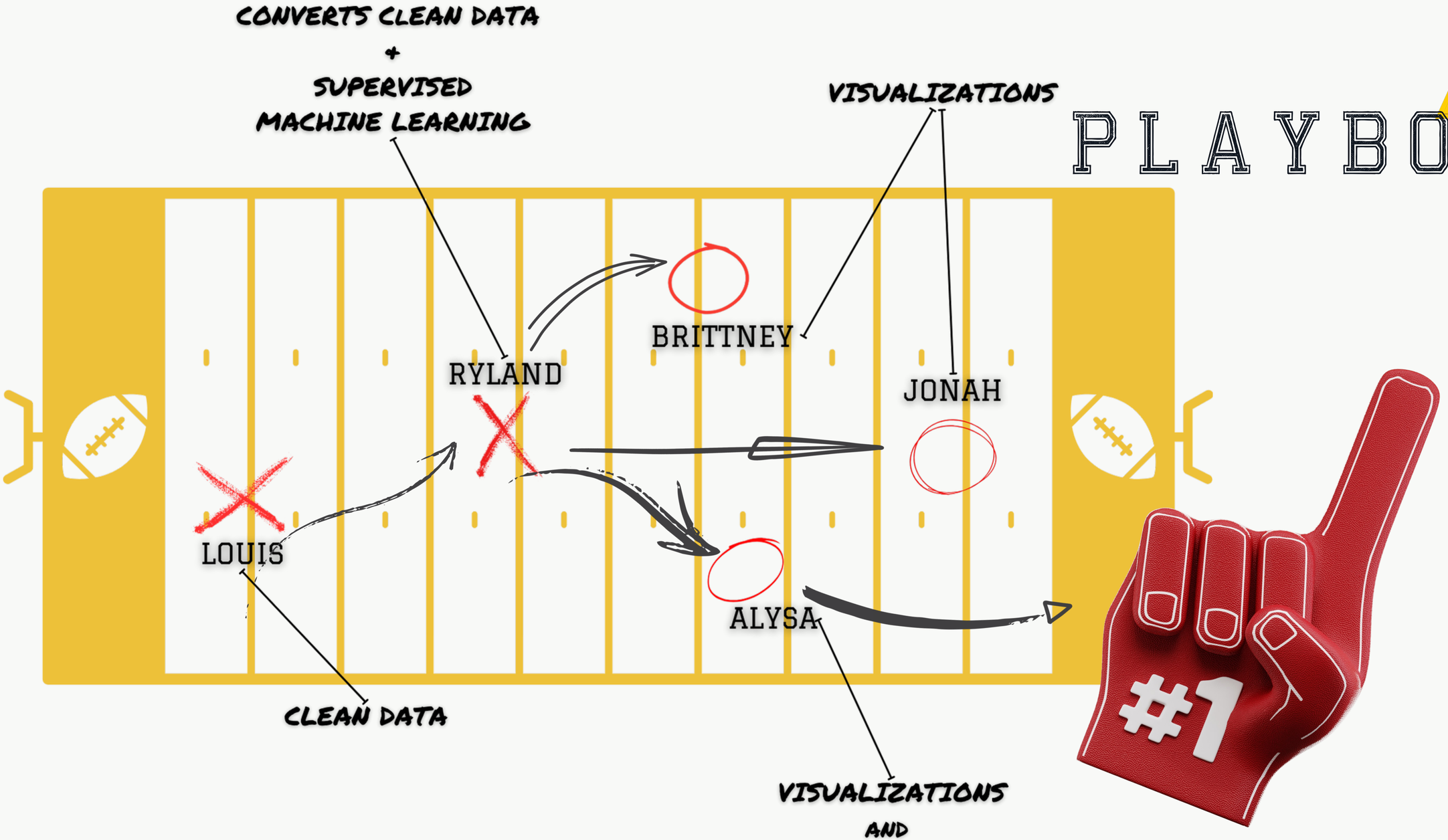
```
gigadata.to_csv("Data/tableau_data.csv")
```

```
gigadata = gigadata.drop(columns =['team_home', 'score_home', 'score_away', 'team_away','stadium'] )
gigadata.columns
```

```
Index(['schedule_date', 'schedule_season', 'schedule_week', 'schedule_playoff',
       'team_favorite_id', 'spread_favorite', 'over_under_line',
       'stadium_neutral', 'weather_temperature', 'weather_wind_mph',
       'weather_humidity', 'weather_detail', 'winner(0_away,1_home,2_tie)',
       'team_name_home', 'team_name_short_home', 'team_id_home',
       'team_id_pfr_home', 'team_conference_home', 'team_division_home',
       'team_conference_pre2002_home', 'team_division_pre2002_home',
       'team_name_away', 'team_name_short_away', 'team_id_away',
       'team_id_pfr_away', 'team_conference_away', 'team_division_away',
       'team_conference_pre2002_away', 'team_division_pre2002_away',
       'stadium_name', 'stadium_location', 'stadium_type', 'stadium_address',
       'stadium_weather_type', 'stadium_surface', 'stadium_latitude',
       'stadium_longitude', 'stadium_elevation'],
      dtype='object')
```

```
gigadata['schedule_date'] = gigadata['schedule_date'].str.replace(r'\D', '', regex=True).astype(int)
gigadata.head()
```

PLAYBOOK

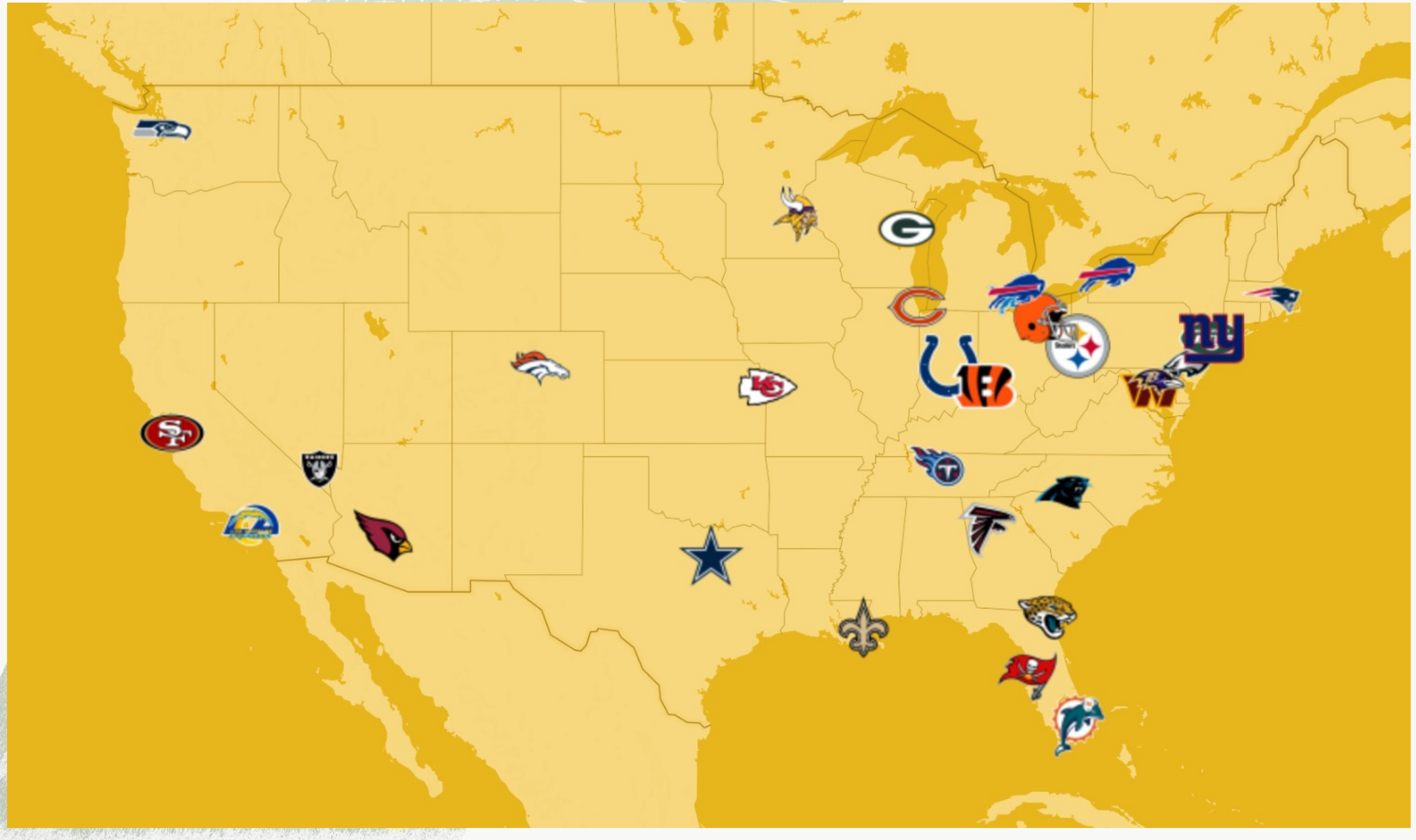

CONVERTS CLEAN DATA
+
SUPERVISED
MACHINE LEARNING
VISUALIZATIONS
BRITTNEY
RYLAND
JONAH
LOUIS
ALYSA
CLEAN DATA
VISUALIZATIONS
AND
PRESENTATION

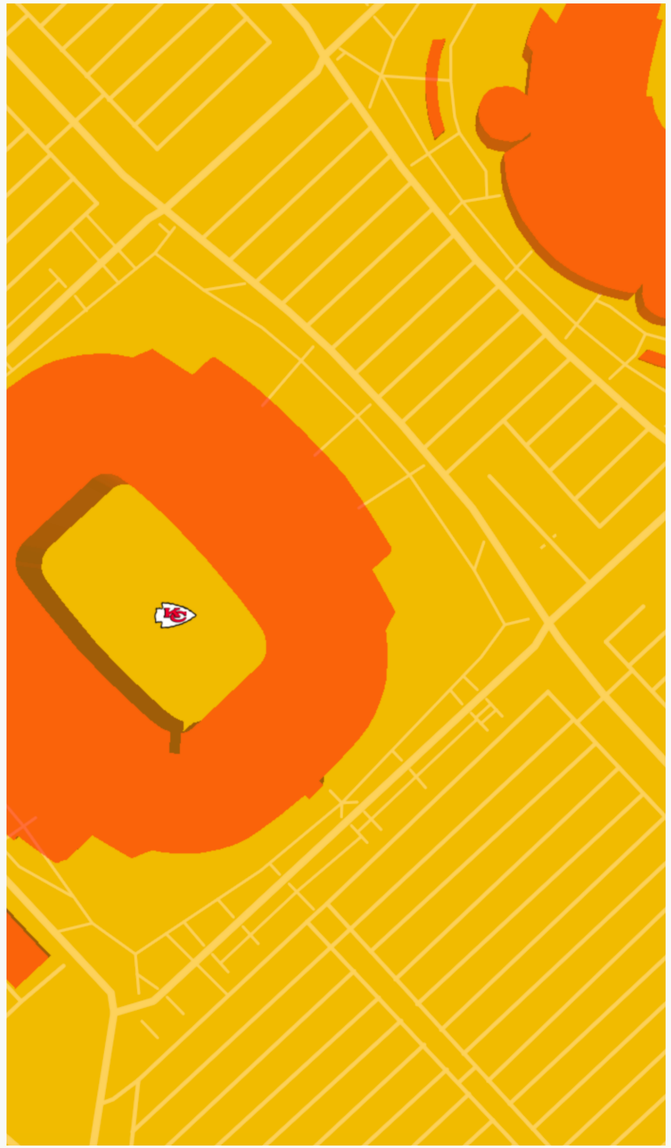
#1

TOUCHDOWNS + trends

QUESTION NO. ONE

WHAT CAN WE LEARN FROM THE HISTORY AND CHANGES OF NFL TEAMS BY LOOKING AT THEIR STADIUMS?





1997

2012

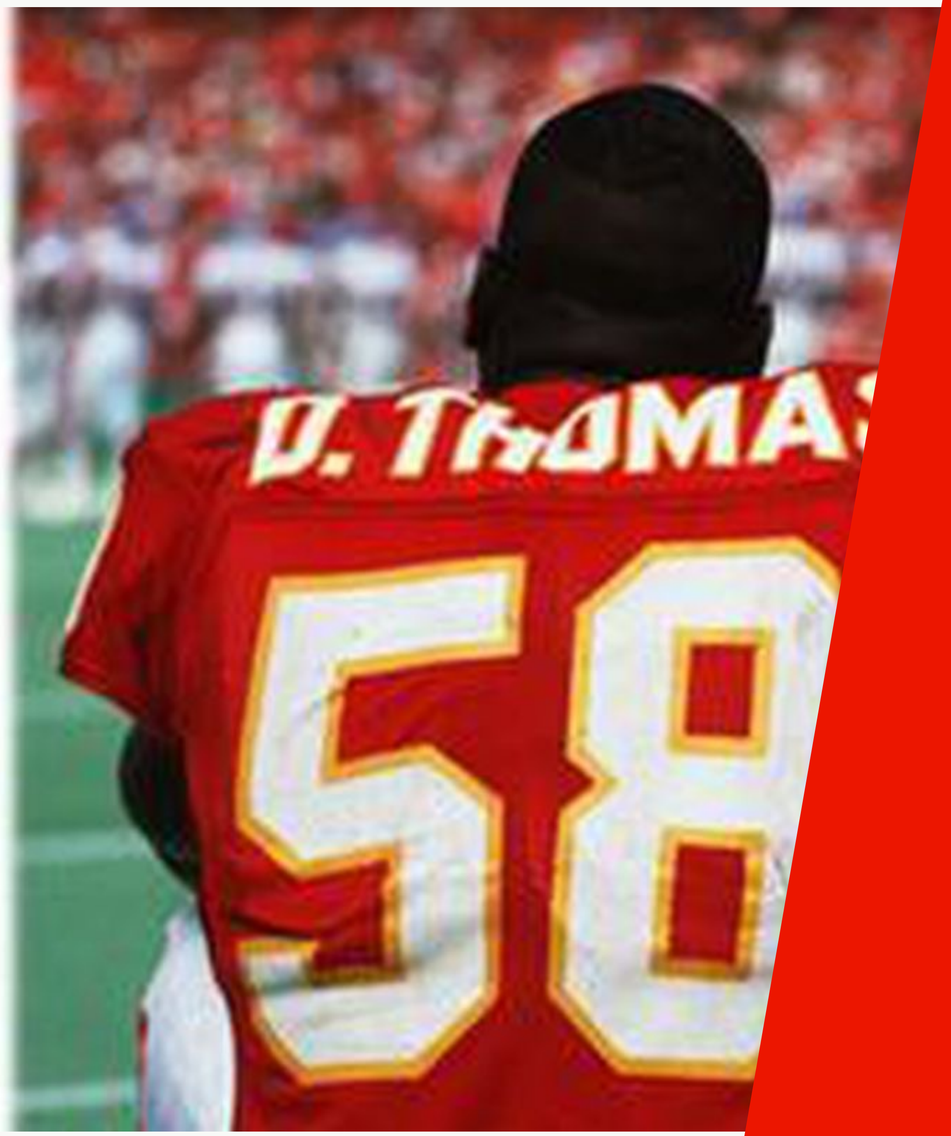
Team Home:	Kansas City Chiefs
Stadium Name:	Arrowhead Stadium
Stadium City:	Kansas City, MO
Stadium Surface:	Grass
Stadium Type:	outdoor
Stadium Weather Type:	cold
Avg. Weather Temperature:	48.75 Degrees (F)
Avg. MPH Winds:	8.00 MPH
Stadium Elevation:	2,119
Avg. Score by Kansas City Chiefs:	32.75
Avg. Score by Opponent:	16.00
Number of Home Games Won:	8

2018

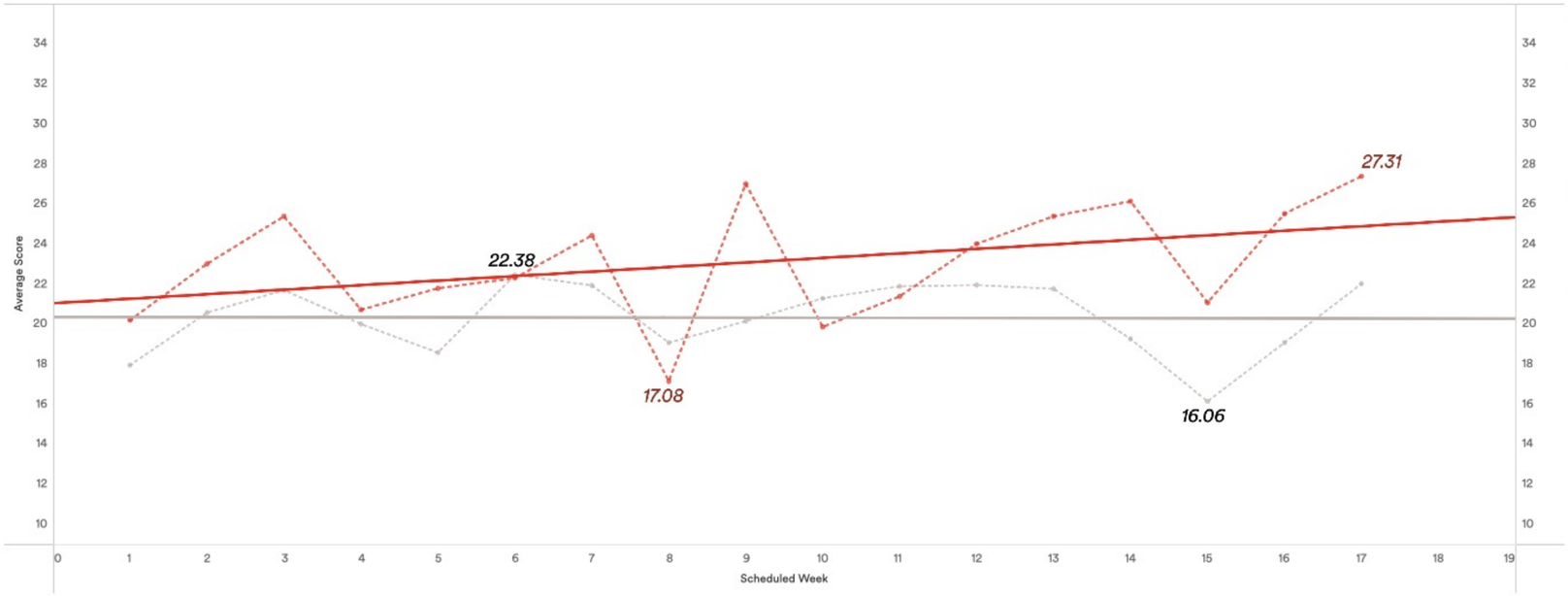
Team Home:	Kansas City Chiefs
Stadium Name:	Arrowhead Stadium
Stadium City:	Kansas City, MO
Stadium Surface:	Grass
Stadium Type:	outdoor
Stadium Weather Type:	cold
Avg. Weather Temperature:	61.00 Degrees (F)
Avg. MPH Winds:	0.00 MPH
Stadium Elevation:	265
Avg. Score by Kansas City Chiefs:	27.00
Avg. Score by Opponent:	21.00
Number of Home Games Won:	1

QUESTION NO. TWO

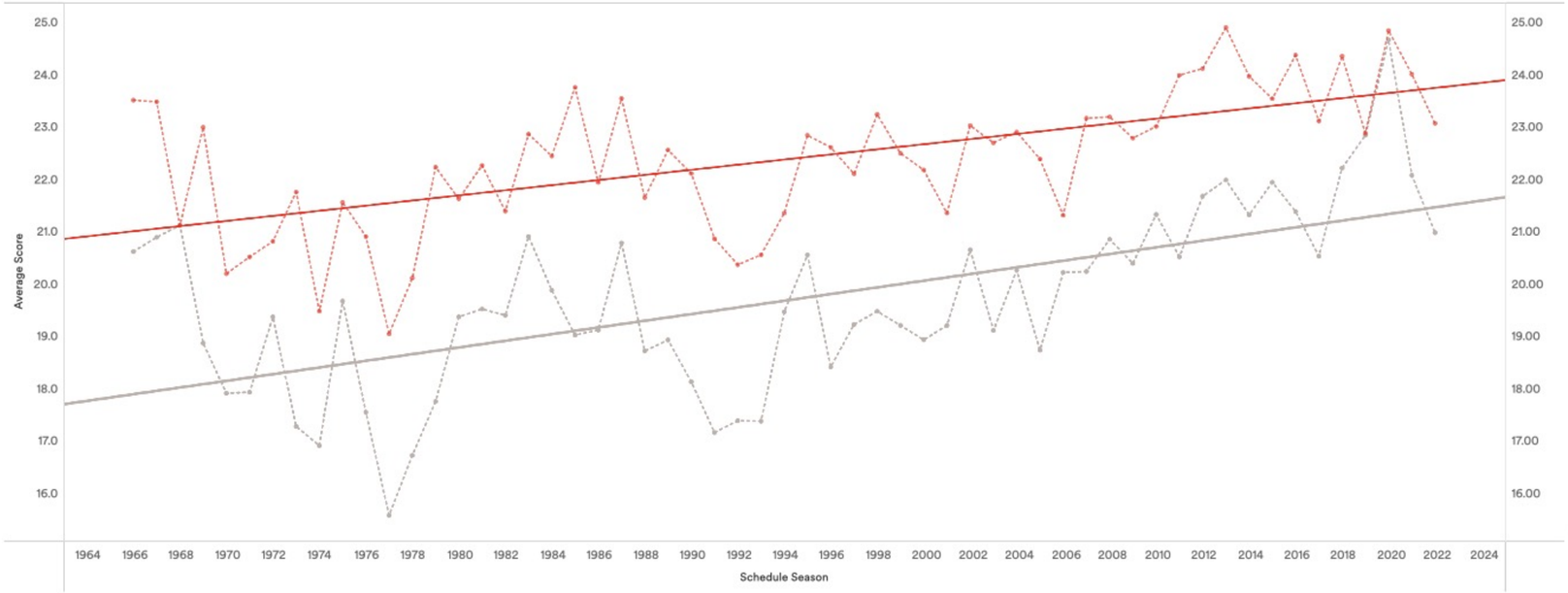
HOW MUCH DOES “HOME FIELD ADVANTAGE” AFFECT SCORING THROUGHOUT SCHEDULED WEEKS? SEASONS?



Scheduled Week Analysis: Average Home vs Away Scores



Season Analysis: Average Home vs Away Scores



Measure Names
- Overall Avg. Away Score
- Overall Avg. Home Score

Overall Avg. Scores

Overall Avg. Home Score	22.5
Overall Avg. Away Score	19.8

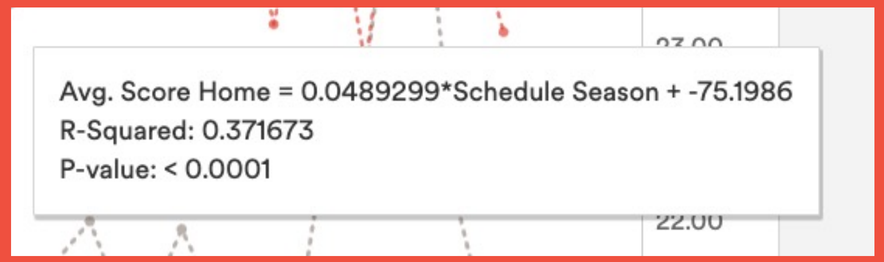
Schedule Season: 2007

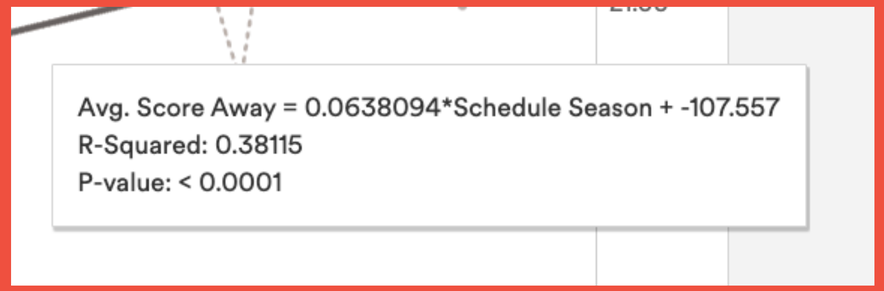
Show history

Seasonal Avg. Score Variation

	Overall..	% Diffe..	Overall..	% Diffe..
1966	23.50		20.61	
1967	23.47	-0.13%	20.88	1.34%
1968	21.11	-10.08%	21.11	1.09%
1969	22.98	8.90%	18.86	-10.68%
1970	20.19	-12.18%	17.90	-5.08%
1971	20.51	1.60%	17.92	0.12%
1972	20.80	1.44%	19.36	8.03%
1973	21.75	4.53%	17.26	-10.82%
1974	19.47	-10.46%	16.90	-2.11%
1975	21.55	10.68%	19.66	16.34%
1976	20.90	-3.03%	17.54	-10.81%
1977	19.03	-8.91%	15.57	-11.24%
1978	20.10	5.61%	16.71	7.36%
1979	22.22	10.55%	17.74	6.16%
1980	21.62	-2.72%	19.36	9.12%
1981	22.26	2.96%	19.51	0.78%
1982	21.38	-3.93%	19.39	-0.62%
1983	22.85	6.88%	20.90	7.77%
1984	22.43	-1.84%	19.87	-4.93%
1985	23.75	5.85%	19.01	-4.30%
1986	21.93	-7.65%	19.11	0.50%
1987	23.54	7.32%	20.77	8.72%
1988	21.64	-8.06%	18.71	-9.94%
1989	22.55	4.20%	18.92	1.12%
1990	22.10	-1.98%	18.12	-4.23%
1991	20.86	-5.64%	17.15	-5.35%
1992	20.36	-2.39%	17.37	1.32%
1993	20.54	0.92%	17.37	-0.05%
1994	21.34	3.89%	19.46	12.03%
1995	22.83	6.97%	20.55	5.61%
1996	22.60	-1.01%	18.40	-10.43%
1997	22.10	-2.24%	19.22	4.42%
1998	23.23	5.14%	19.47	1.31%
1999	22.49	-3.21%	19.20	-1.38%
2000	22.17	-1.41%	18.92	-1.43%
2001	21.34	-3.73%	19.20	1.45%
2002	23.02	7.85%	20.64	7.52%
2003	22.69	-1.43%	19.10	-7.46%
2004	22.90	0.91%	20.25	6.02%
2005	22.38	-2.26%	18.72	-7.55%
2006	21.30	-4.80%	20.21	7.94%
2007	23.16	8.70%	20.23	0.09%
2008	23.18	0.11%	20.84	3.04%
2009	22.78	-1.74%	20.38	-2.21%
2010	23.00	0.97%	21.32	4.59%

Measure Values
-24.89 24.89





- Annual increase: 0.049 points
- One-point increase: ~20.4 years
- Linear relationship explains 37.17% variability in home team scores
- P-value < 0.0001: Strong statistical association with schedule season

HOME

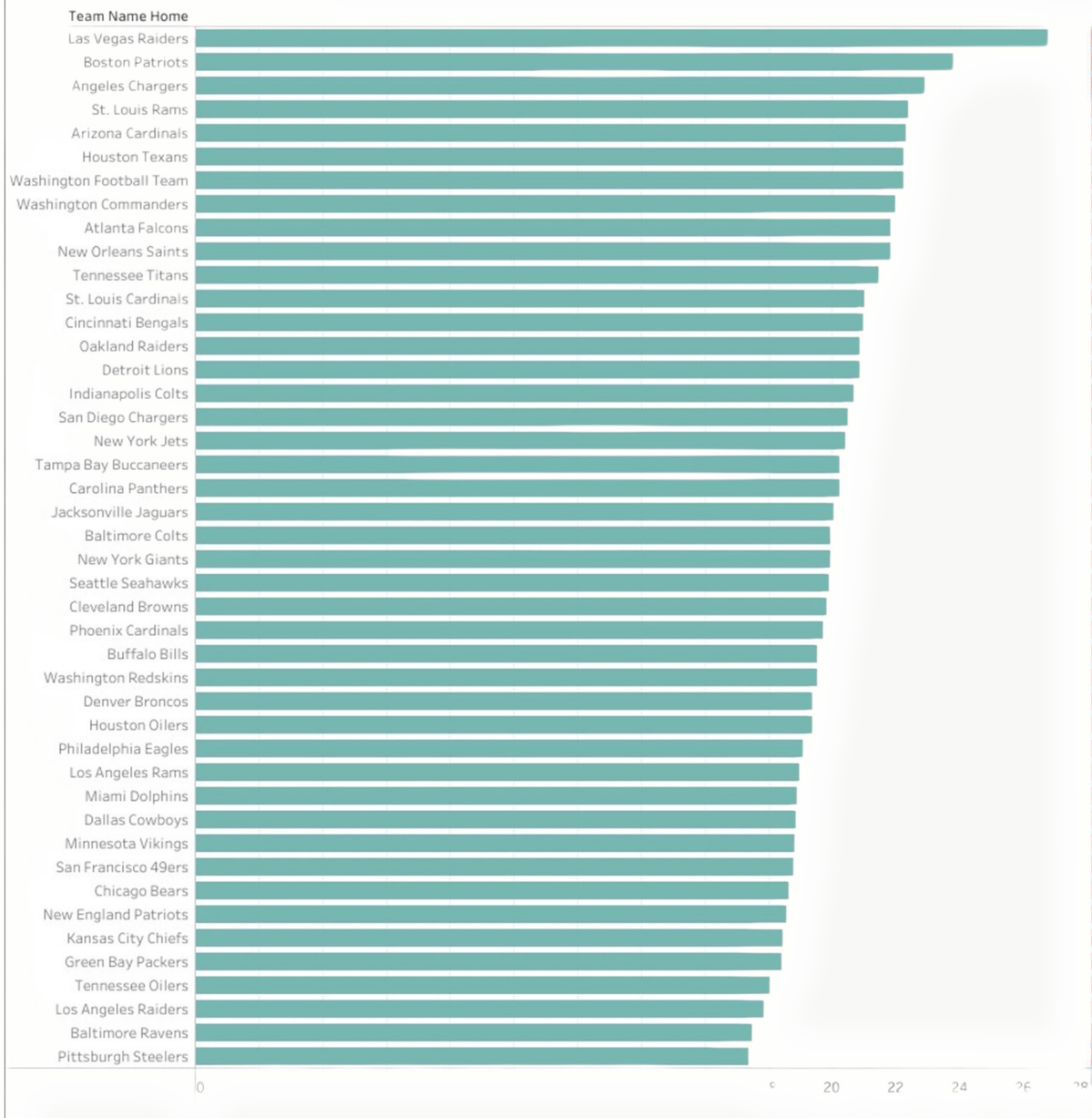
- Annual increase: 0.064 points
- One-point increase: ~15.7 years
- Linear relationship explains 38.12% variability in home team scores
- P-value < 0.0001: Strong statistical association with schedule season

AWAY

QUESTION NO. 3

CAN THE GAME'S OUTCOME BE AFFECTED BY WHETHER A TEAM PLAYS AT HOME OR AWAY?

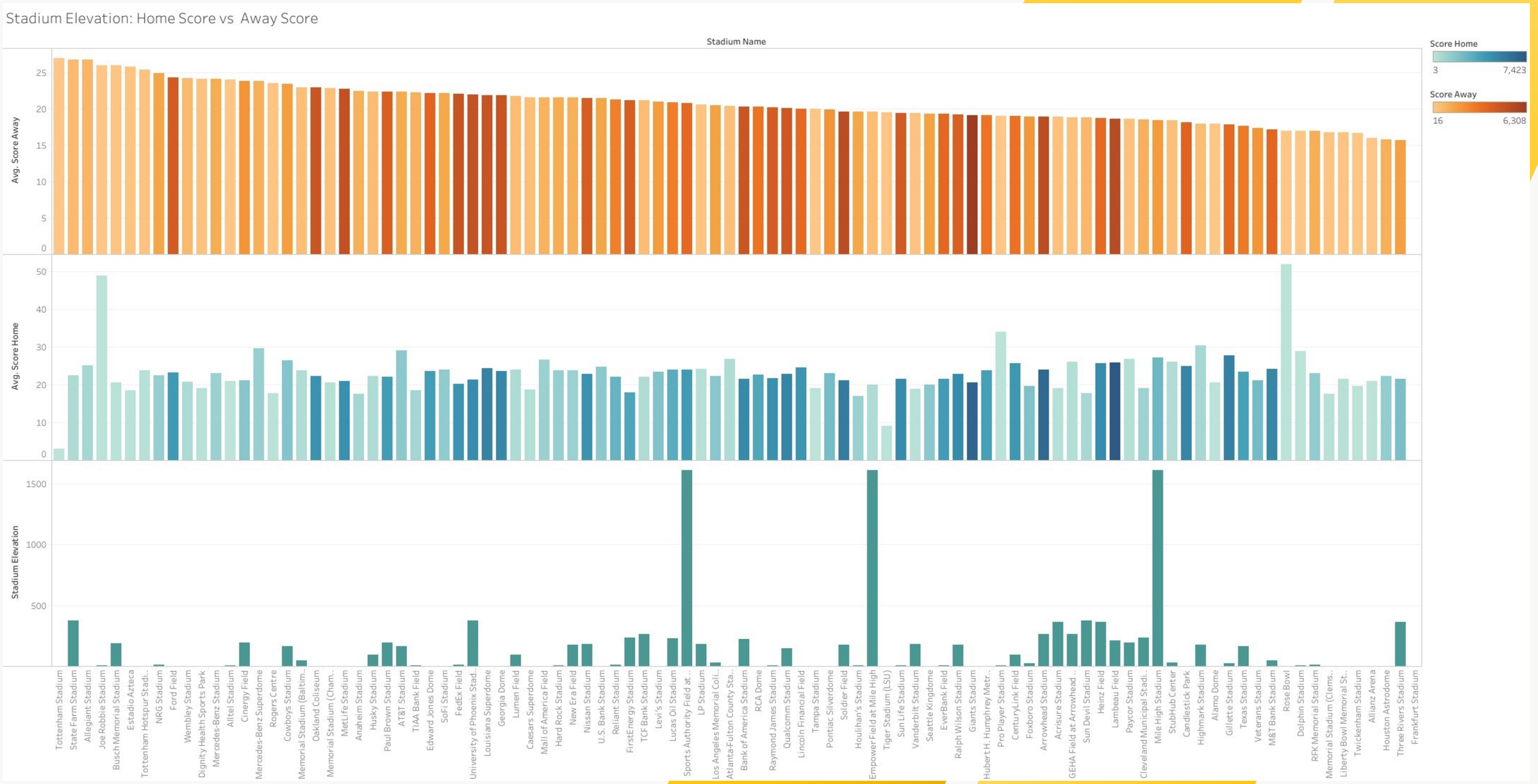
Team Home vs. Team Away



QUESTION NO. FOUR

DOES STADIUM ELEVATION AFFECT WHO WILL WIN THE FOOTBALL GAME?

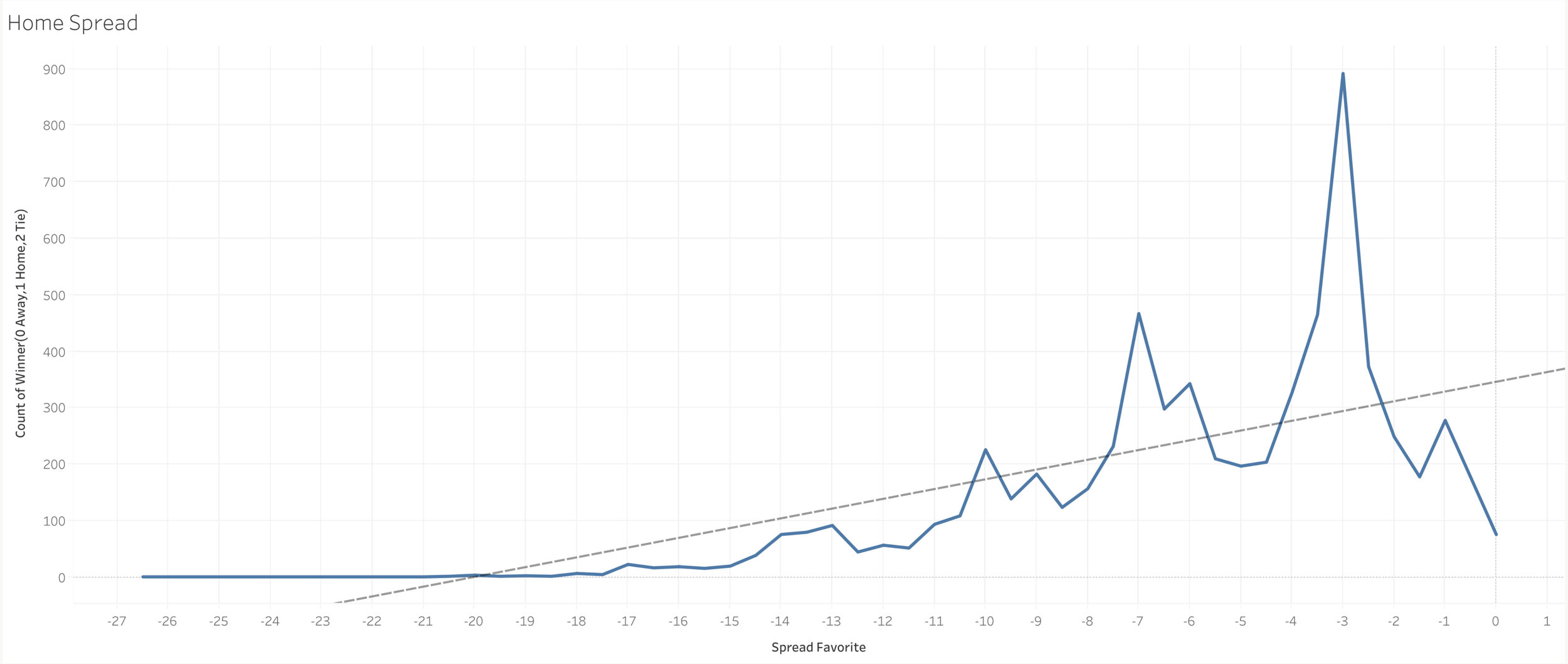
Stadium Elevation: Home Score vs Away Score



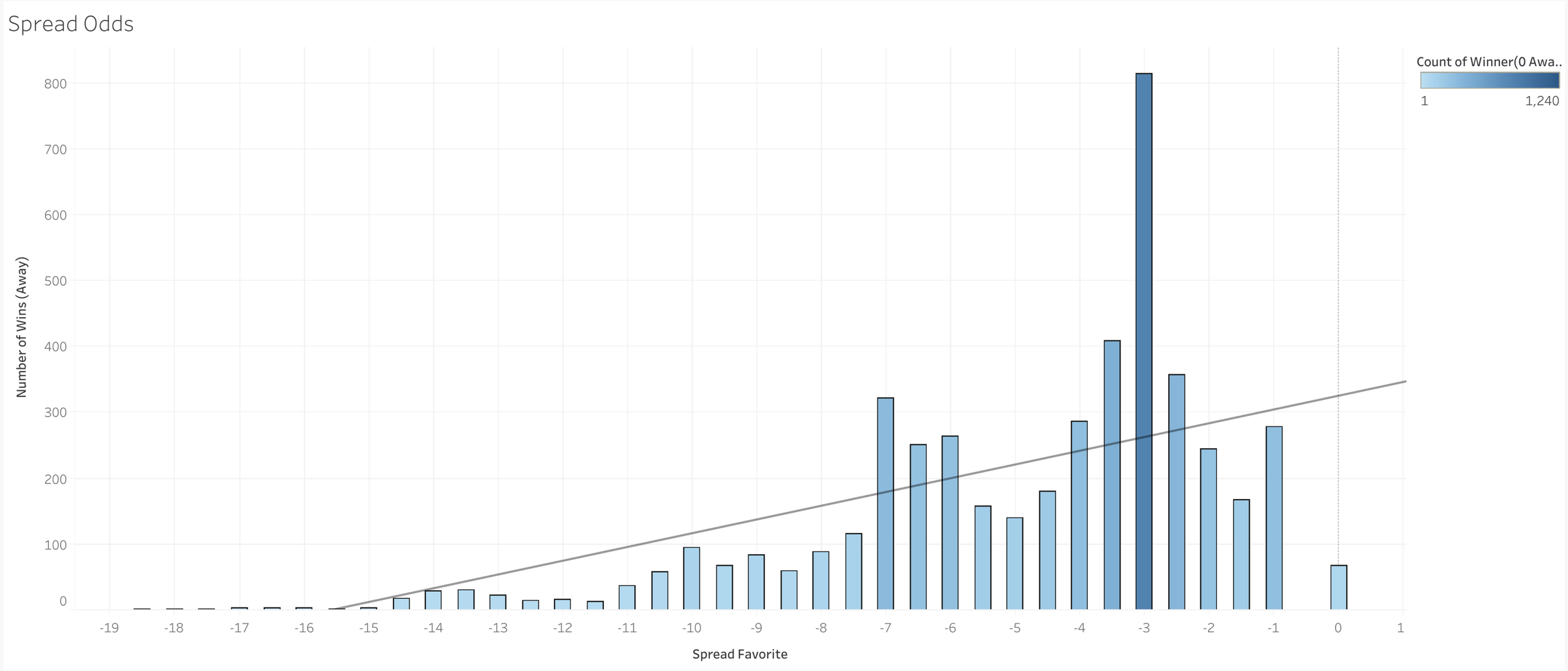
QUESTION NO. FOUR

HOW DOES SPREAD AFFECT WHO WILL WIN THE GAME?

HOME SPREAD


Home Spread
Count of Winner(0 Away,1 Home,2 Tie)
900
800
700
600
500
400
300
200
100
0
-27
-26
-25
-24
-23
-22
-21
-20
-19
-18
-17
-16
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
Spread Favorite

AWAY SPREAD


Spread Odds
Count of Winner(0 Awa..
1
1,240
Number of Wins (Away)
800
700
600
500
400
300
200
100
0
-19
-18
-17
-16
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
Spread Favorite

MACHINE LEARNING

MOCKUP

SUPERVISED MACHINE LEARNING

In [10]:
```
tuner.search(X_train_scaled,y_train,epochs=20,validation_data=(X_test_scaled,y_test))
```

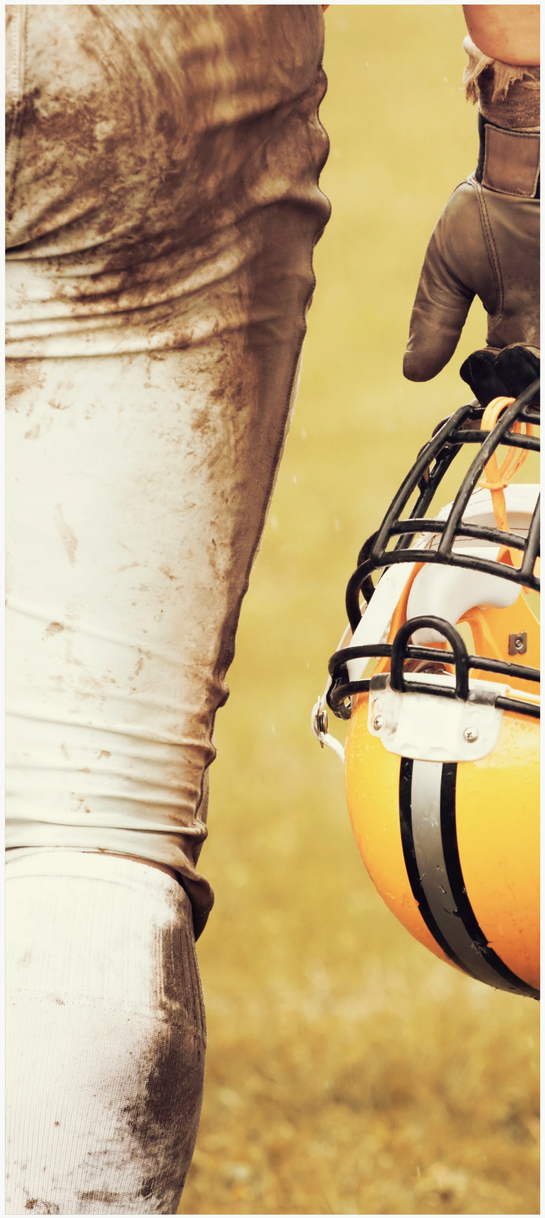
```
Trial 60 Complete [00h 00m 22s]
val_accuracy: 0.5720916986465454

Best val_accuracy So Far: 0.6106759309768677
Total elapsed time: 00h 10m 40s
```

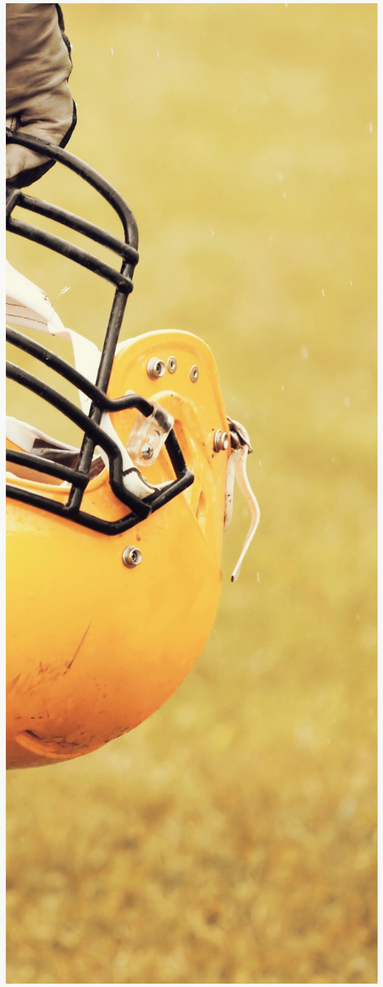
In [11]:
```
# Get best model hyperparameters
best_hyper = tuner.get_best_hyperparameters(1)[0]
best_hyper.values
```

Out[11]:
```
{'activation': 'tanh',
 'first_units': 9,
 'num_layers': 2,
 'units_0': 5,
 'units_1': 3,
 'units_2': 5,
 'units_3': 3,
 'units_4': 7,
 'units_5': 3,
 'tuner/epochs': 20,
 'tuner/initial_epoch': 0,
 'tuner/bracket': 0,
 'tuner/round': 0}
```

In [12]:
```
# Evaluate best model against full test data
best_model = tuner.get_best_models(1)[0]
model_loss, model_accuracy = best_model.evaluate(X_test_scaled,y_test,verbose=2)
print(f"Loss: {model_loss}, Accuracy: {model_accuracy}")
```

```
108/108 - 0s - loss: 0.6492 - accuracy: 0.6107 - 347ms/epoch - 3ms/step
Loss: 0.6491859555244446, Accuracy: 0.6106759309768677
```





MACHINE LEARNING

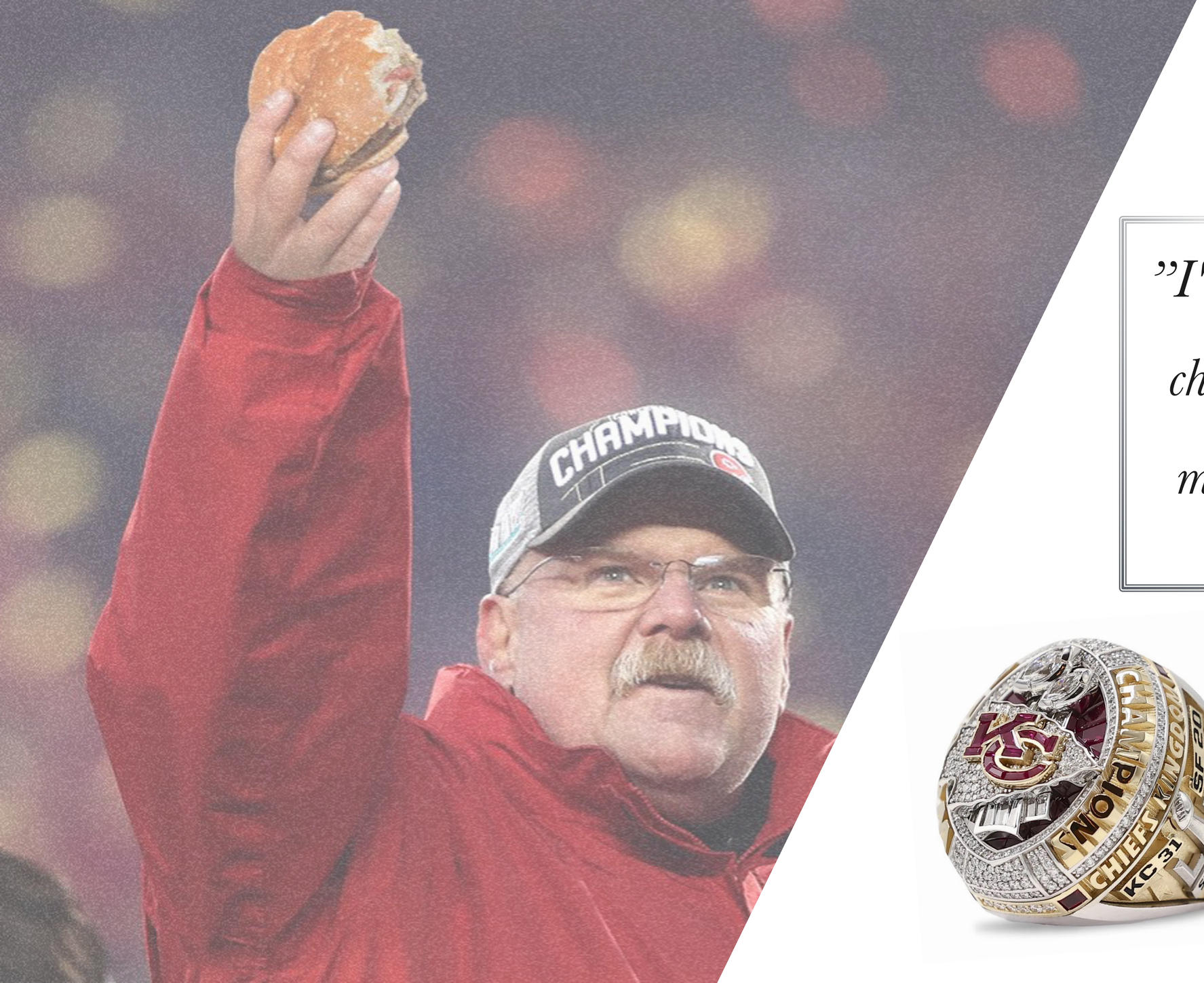
RESULTS

LOSS

64.9%

ACCURACY

61.1%



”I'm going to get the biggest cheeseburger you can find, might make it a double.”

-ANDY REID-

SUPERBOWL LIV RING
2020


THANK YOU
thank you
so much!


ANY
QUESTIONS
?