Super Bowl Offensive Statistics 1966-2021
Exploratory Analysis

Sam Jones, sjones10@bellarmine.edu
Cason Henderson, chenderson2@bellarmine.edu

I.   INTRODUCTION

As a group, we choose to do our data on offensive Super Bowl stats. This included Superbowl Offensive Stats from 1966-2021. This data had 23 columns and 729 rows. Passing and rushing stats were the most important, and Team and player names least important. We found our data on Kaggle. The reason we chose to use this data was because we are avid sports fans and wanted to compare past Super Bowl stats to see in future games.

II.   SUPER BOWL DATA SET DESCRIPTION

This data set contains 729 rows with 23 columns with various data types. A listing of data types and missing data is in **Table 1**. For data types, all our missing data were nominal and float numbers that we had to correct for the Pandas data type. For example, the qbr was nominal/float64. For the missing data we calculated it by doing the number of missing entries divided by what the total sum was supposed to be.

Table 1: Data Types and Missing Data

Variable Name	Data Type	Missing Data (%)
V1 – times sacked	Nominal, float	49%
V2 – sack yards	Nominal, float	45%
V3 – qbr	Nominal, float	85%

III.   Super Bowl Data Set Summary Statistics

For this section we will be displaying tables of our data set for Super Bowl Offensive Data. By using functions like .describe, .value_counts, and .corr we were able to show a variety of results and patterns.

Table 2: Summary Statistics for Super Bowl Offensive Data

Variable Name	Count	Mean	Standard Deviation	Min	25th	50th	75th	Max
Year	1019	1994	15.75	1966	1981	1994	2008	2021
Completions	1019	2.16	6.48	0	0	0	0	43
Pass Attempts	1019	3.66	10.71	0	0	0	0	62
Passing Yards	1019	26.75	79.24	0	0	0	0	505
Passing Touchdowns	1019	0.16	0.61	0	0	0	0	6
Interceptions	1019	0.13	0.54	0	0	0	0	5
Longest Pass	1019	0.56	1.28	0	0	0	0	7
Rushing Attempts	1019	3.64	9.46	0	0	0	3	64
Rushing Yards	1019	4.89	14.15	-10	0	0	13	85
Rushing Touchdowns	1019	3.11	5.79	0	0	0	0	38
Longest Run	1019	12.58	26.48	-10	0	0	7	75

Receptions	1019	2.15	2.14	-8	0	1	3	14
Receiving Yards	1019	26.76	34.11	-7	0	14	42	215
Receiving Touchdowns	1019	0.15	0.49	-8	0	0	0	3
Longest Reception	1019	13.25	14.93	-7	0	10	20	85
Fumbles	1019	0.13	0.39	0	0	0	0	3

Table 3: Proportions for Super Bowl Offensive Data

For this table we used the highest frequent and proportion for categorical variables since the code is in descending order.

Category	Frequency	Proportion (%)
Year - 1989	25	2.45%
Completions – 1	13	1.27%
Pass Attempts – 1	18	1.76%
Passing Yards - 256	3	0.29%
Passing Touchdowns - 1	45	4.41%
Interceptions - 1	41	4.02%
Longest Pass - 27	7	0.68%
Rushing Attempts - 1	137	13.44%
Rushing Yards - 1	22	2.15%
Rushing Touchdowns - 1	71	6.96%
Longest Run – 5	31	3.04%
Receptions - 1	211	20.70%
Receiving Yards - 7	22	2.15%
Receiving Touchdowns - 1	129	12.65%
Longest Reception – 7	37	3.63%
Fumbles - 1	107	10.50%

Table 4: Correlation Table/Tables

```
              id         year     completions    pass_attempts
id          1.000000  -0.999335   -0.056734      -0.040972
year       -0.999335   1.000000    0.051692       0.035741
completions -0.056734  0.051692    1.000000       0.985385
pass_attempts -0.040972 0.035741   0.985385       1.000000
passing_yards -0.039226 0.034494   0.969260       0.964177
passing_td  -0.032338   0.029533   0.780092       0.746753
interception 0.037000  -0.041437   0.639645       0.705750
times_sacked -0.002411  0.002028   0.720462       0.761742
sack_yards   0.012767  -0.013986   0.649034       0.705401
longest_pass 0.038110  -0.043195   0.817886       0.835513
qbr         -0.130397   0.134742   0.225309       0.044540
rush_attempts 0.093713 -0.097366  -0.043687      -0.045031
```

rushing_yards	0.061604	-0.065231	-0.028610	-0.028047
rushing_td	0.033746	-0.035505	0.016834	0.016057
longest_run	0.046882	-0.051509	0.042591	0.042635
receptions	-0.140246	0.139769	-0.309055	-0.319074
receiving_yards	-0.079456	0.080077	-0.260090	-0.264818
receiving_td	-0.037285	0.039426	-0.165244	-0.185355
longest_reception	-0.000840	0.002352	-0.291362	-0.295934
fumble	0.054824	-0.057147	0.295978	0.320754
fumbles_lost	0.541308	-0.542487	0.208681	0.232850

	passing_yards	passing_td	interception	times_sacked
id	-0.039226	-0.032338	0.037000	-0.002411
year	0.034494	0.029533	-0.041437	0.002028
completions	0.969260	0.780092	0.639645	0.720462
pass_attempts	0.964177	0.746753	0.705750	0.761742
passing_yards	1.000000	0.821743	0.607054	0.724354
passing_td	0.821743	1.000000	0.412350	0.517271
interception	0.607054	0.412350	1.000000	0.583502
times_sacked	0.724354	0.517271	0.583502	1.000000
sack_yards	0.653812	0.432656	0.567638	0.943434
longest_pass	0.894395	0.754331	0.574812	0.685294
qbr	0.352480	0.600792	-0.519581	-0.064331
rush_attempts	-0.040943	-0.027469	-0.037463	-0.075660
rushing_yards	-0.024355	-0.009963	-0.032480	-0.031142
rushing_td	0.033359	-0.013323	0.005611	0.018163
longest_run	0.049731	0.059309	0.012621	0.055125
receptions	-0.308098	-0.242635	-0.246698	-0.362356
receiving_yards	-0.261738	-0.209859	-0.194225	-0.303945
receiving_td	-0.152360	-0.127738	-0.192656	-0.216946
longest_reception	-0.292385	-0.233835	-0.213974	-0.333678
fumble	0.304200	0.216157	0.299775	0.404143
fumbles_lost	0.228613	0.167773	0.173588	0.326619

	sack_yards	longest_pass		rush_attempts
id	0.012767	0.038110	...	0.093713
year	-0.013986	-0.043195	...	-0.097366
completions	0.649034	0.817886	...	-0.043687
pass_attempts	0.705401	0.835513	...	-0.045031
passing_yards	0.653812	0.894395	...	-0.040943
passing_td	0.432656	0.754331	...	-0.027469
interception	0.567638	0.574812	...	-0.037463
times_sacked	0.943434	0.685294	...	-0.075660
sack_yards	1.000000	0.637086	...	-0.067851
longest_pass	0.637086	1.000000	...	-0.038235
qbr	-0.138299	0.378810	...	0.233142
rush_attempts	-0.067851	-0.038235	...	1.000000
rushing_yards	-0.027812	-0.029342	...	0.908461
rushing_td	-0.002969	0.036024	...	0.557756
longest_run	0.046536	0.045836	...	0.682194
receptions	-0.336401	-0.304640	...	-0.052451
receiving_yards	-0.275412	-0.263104	...	-0.155316

receiving_td	-0.228529	-0.149889	...	-0.108458
longest_reception	-0.297392	-0.291634	...	-0.146028
fumble	0.373496	0.300210	...	0.133407
fumbles_lost	0.351335	0.306839	...	0.164297

	rushing_yards	rushing_td	longest_run	receptions
id	0.061604	0.033746	0.046882	-0.140246
year	-0.065231	-0.035505	-0.051509	0.139769
completions	-0.028610	0.016834	0.042591	-0.309055
pass_attempts	-0.028047	0.016057	0.042635	-0.319074
passing_yards	-0.024355	0.033359	0.049731	-0.308098
passing_td	-0.009963	-0.013323	0.059309	-0.242635
interception	-0.032480	0.005611	0.012621	-0.246698
times_sacked	-0.031142	0.018163	0.055125	-0.362356
sack_yards	-0.027812	-0.002969	0.046536	-0.336401
longest_pass	-0.029342	0.036024	0.045836	-0.304640
qbr	0.165885	0.011040	0.156495	0.060161
rush_attempts	0.908461	0.557756	0.682194	-0.052451
rushing_yards	1.000000	0.543708	0.847226	-0.043151
rushing_td	0.543708	1.000000	0.437550	-0.025502
longest_run	0.847226	0.437550	1.000000	-0.076743
receptions	-0.043151	-0.025502	-0.076743	1.000000
receiving_yards	-0.134359	-0.078286	-0.161189	0.847410
receiving_td	-0.097766	-0.051042	-0.097823	0.393834
longest_reception	-0.120936	-0.068412	-0.152796	0.573804
fumble	0.100741	0.076380	0.078652	-0.071705
fumbles_lost	0.085207	0.036366	0.012075	-0.101825

	receiving_yards	receiving_td	longest_reception	fumble
id	-0.079456	-0.037285	-0.000840	0.054824
year	0.080077	0.039426	0.002352	-0.057147
completions	-0.260090	-0.165244	-0.291362	0.295978
pass_attempts	-0.264818	-0.185355	-0.295934	0.320754
passing_yards	-0.261738	-0.152360	-0.292385	0.304200
passing_td	-0.209859	-0.127738	-0.233835	0.216157
interception	-0.194225	-0.192656	-0.213974	0.299775
times_sacked	-0.303945	-0.216946	-0.333678	0.404143
sack_yards	-0.275412	-0.228529	-0.297392	0.373496
longest_pass	-0.263104	-0.149889	-0.291634	0.300210
qbr	-0.106111	0.065664	-0.035920	-0.124363
rush_attempts	-0.155316	-0.108458	-0.146028	0.133407
rushing_yards	-0.134359	-0.097766	-0.120936	0.100741
rushing_td	-0.078286	-0.051042	-0.068412	0.076380
longest_run	-0.161189	-0.097823	-0.152796	0.078652
receptions	0.847410	0.393834	0.573804	-0.071705
receiving_yards	1.000000	0.433792	0.836539	-0.062134
receiving_td	0.433792	1.000000	0.400707	-0.100208
longest_reception	0.836539	0.400707	1.000000	-0.085281
fumble	-0.062134	-0.100208	-0.085281	1.000000

```
fumbles_lost              -0.107046     -0.081549         -0.143421  0.675111
```

```
                  fumbles_lost
id                    0.541308
year                 -0.542487
completions           0.208681
pass_attempts         0.232850
passing_yards         0.228613
passing_td            0.167773
interception          0.173588
times_sacked          0.326619
sack_yards            0.351335
longest_pass          0.306839
qbr                  -0.053499
rush_attempts         0.164297
rushing_yards         0.085207
rushing_td            0.036366
longest_run           0.012075
receptions           -0.101825
receiving_yards      -0.107046
receiving_td         -0.081549
longest_reception    -0.143421
fumble                0.675111
fumbles_lost          1.000000
```
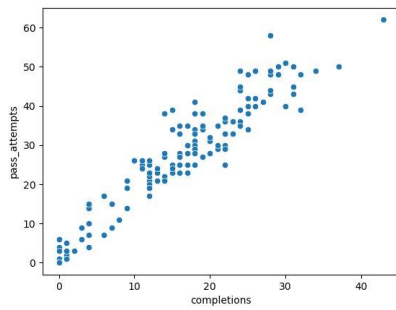
IV.　　　DATA SET GRAPHICAL EXPLORATION

A. *Distributions*
- Not very useful
- Uneven player distribution
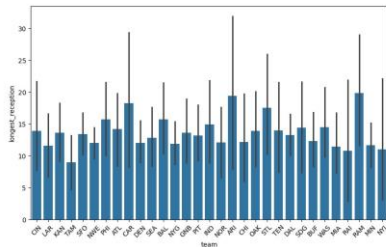- Helps moving forward

Distribution of fumble
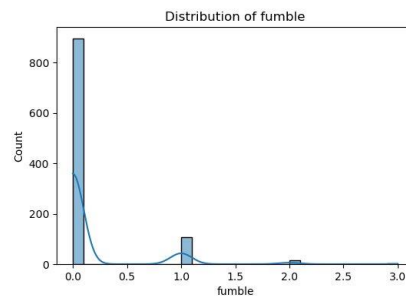
B. *ScatterPlots / Pairwise Plots*
- Very helpful for majority numerical data
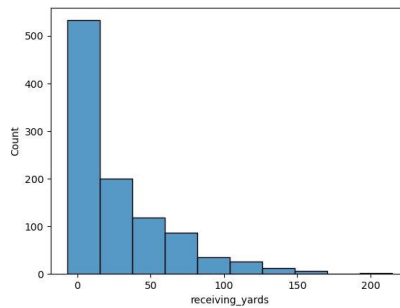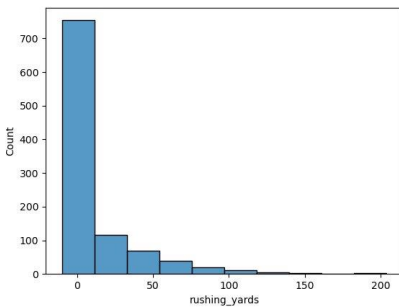- Display patterns and direct relationships
- Highlights outliers
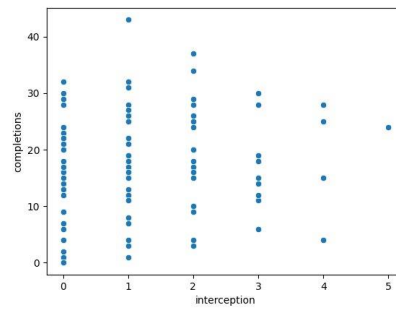
C. *Barcharts (categorical variables)*
- Able to implement categorical data
- Compare categorical to numerical
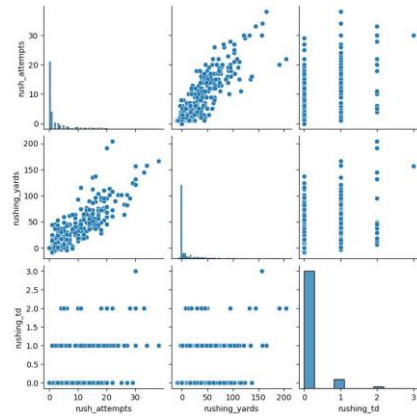- Easiest graph to understand Supports passing data

D. *Histogram*
- Disproportionate
- Not very useful for our focus
- Similar to distribution graphs

E. Bivariate Analysis –Pairplot
- Compares 3 variables together
- Overview of rushing statistics
- Scatter>Bar

V.   SUMMARY OF FINDINGS

 We came to conclusion that defense plays very well in the super bowl. We only found one QB that did really well in the super bowl and threw for 6 passing touchdowns which was our only outlier. We also have found that in the super bowl, that the passing game was a lot more involved than the rushing game. We also were surprised to see that there were higher passing touchdowns than rushing touchdowns in the super bowl.