

# Exploratory Data Analysis

*Kevin Thompson, Sean Kennedy, Sachin Chavan*

*September 29, 2019*

## 1.1: Load data

```
load("../analysis/data/.RData")
load("../analysis/data/RData.dms")
```

## 1.2: Structures QBCrossSectional and QBPanels

### 1.2.1 QB (Quarterback)

```
str(QBCrossSectional)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame': 453 obs. of 21 variables:
## $ PlayerID : num 6739 4314 13320 18857 14536 ...
## $ Week : num 13 3 10 5 8 4 3 4 8 15 ...
## $ Position : chr "QB" "QB" "QB" "QB" ...
## $ Opponent : chr "NYJ" "HOU" "MIA" "KC" ...
## $ TeamIsHome : logi FALSE TRUE TRUE TRUE TRUE TRUE ...
## $ GameDate : Date, format: "2017-12-03" "2017-09-24" ...
## $ PassingCompletions : num 19 25 21 16 26 25 29 22 19 20 ...
## $ PassingAttempts : num 33 35 35 31 41 34 49 29 30 31 ...
## $ PassingCompletionPercentage: num 57.6 71.4 60 51.6 63.4 73.5 59.2 75.9 63.3 64.5 ...
## $ PassingYards : num 366 378 254 261 452 283 373 316 402 242 ...
## $ PassingYardsPerAttempt : num 11.1 10.8 7.3 8.4 11 8.3 7.6 10.9 13.4 7.8 ...
## $ PassingTouchdowns : num 4 5 4 5 4 4 4 3 4 4 ...
## $ PassingInterceptions : num 0 0 0 0 1 1 0 1 3 0 ...
## $ PassingRating : num 136 146 120 120 123 ...
## $ RushingAttempts : num 1 1 5 3 4 4 7 8 8 14 ...
## $ RushingYards : num 70 6 95 31 30 24 26 44 67 58 ...
## $ RushingYardsPerAttempt : num 70 6 19 10.3 7.5 6 3.7 5.5 8.4 4.1 ...
## $ RushingTouchdowns : num 0 0 0 0 0 1 0 1 0 0 ...
## $ FumblesLost : num 0 1 0 0 0 0 0 0 0 0 ...
## $ FantasyPoints : num 37.6 35.7 35.7 35.5 35.1 ...
## $ Team : chr "KC" "NE" "CAR" "HOU" ...
```

### 1.2.2 QB (Panel Data)

```
str(QBPanels)
```

```
## List of 17
## $ Opponent :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of 18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1 : chr [1:56] NA "CHI" "DET" "SEA" ...
## ..$ 2 : chr [1:56] "CLE" "GB" "IND" "ATL" ...
## ..$ 3 : chr [1:56] NA "DET" "DAL" "CIN" ...
## ..$ 4 : chr [1:56] "PIT" "BUF" "SF" "CHI" ...
## ..$ 5 : chr [1:56] "OAK" NA "PHI" "DAL" ...
## ..$ 6 : chr [1:56] NA "MIA" "TB" NA ...
```

```

## ..$ 7      : chr [1:56] "MIN" "NE" NA NA ...
## ..$ 8      : chr [1:56] "MIA" "NYJ" NA NA ...
## ..$ 9      : chr [1:56] "TEN" "CAR" NA NA ...
## ..$ 10     : chr [1:56] NA "DAL" NA NA ...
## ..$ 11     : chr [1:56] "GB" "SEA" NA NA ...
## ..$ 12     : chr [1:56] "HOU" "TB" NA NA ...
## ..$ 13     : chr [1:56] "DET" NA NA NA ...
## ..$ 14     : chr [1:56] "PIT" NA NA NA ...
## ..$ 15     : chr [1:56] "CLE" "TB" NA "CAR" ...
## ..$ 16     : chr [1:56] "IND" "NO" NA NA ...
## ..$ 17     : chr [1:56] "CIN" "CAR" NA NA ...
## $ TeamIsHome      :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : logi [1:56] NA FALSE FALSE TRUE FALSE TRUE ...
## ..$ 2      : logi [1:56] TRUE TRUE FALSE FALSE TRUE FALSE ...
## ..$ 3      : logi [1:56] NA FALSE TRUE TRUE FALSE TRUE ...
## ..$ 4      : logi [1:56] TRUE TRUE TRUE TRUE FALSE TRUE ...
## ..$ 5      : logi [1:56] FALSE NA FALSE FALSE NA FALSE ...
## ..$ 6      : logi [1:56] NA TRUE TRUE NA FALSE FALSE ...
## ..$ 7      : logi [1:56] FALSE FALSE NA NA TRUE TRUE ...
## ..$ 8      : logi [1:56] TRUE FALSE NA NA FALSE TRUE ...
## ..$ 9      : logi [1:56] FALSE FALSE NA NA NA NA ...
## ..$ 10     : logi [1:56] NA TRUE NA NA FALSE FALSE ...
## ..$ 11     : logi [1:56] FALSE FALSE NA NA TRUE FALSE ...
## ..$ 12     : logi [1:56] TRUE TRUE NA NA TRUE TRUE ...
## ..$ 13     : logi [1:56] TRUE NA NA NA FALSE FALSE ...
## ..$ 14     : logi [1:56] FALSE NA NA NA TRUE FALSE ...
## ..$ 15     : logi [1:56] FALSE FALSE NA FALSE TRUE FALSE ...
## ..$ 16     : logi [1:56] TRUE FALSE NA NA FALSE TRUE ...
## ..$ 17     : logi [1:56] TRUE TRUE NA NA NA TRUE ...
## $ PassingCompletions :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 21 27 28 24 16 NA NA 26 NA ...
## ..$ 2      : num [1:56] 25 19 19 33 23 30 NA 22 16 NA ...
## ..$ 3      : num [1:56] NA 24 29 28 22 25 NA 35 18 NA ...
## ..$ 4      : num [1:56] 31 24 33 18 18 32 NA 30 NA NA ...
## ..$ 5      : num [1:56] 19 NA 28 19 NA 30 NA 21 23 NA ...
## ..$ 6      : num [1:56] NA 24 18 NA 17 20 NA 11 31 NA ...
## ..$ 7      : num [1:56] 27 23 NA NA 14 21 NA 19 17 13 ...
## ..$ 8      : num [1:56] 10 18 NA NA 17 32 NA NA 26 NA ...
## ..$ 9      : num [1:56] 34 24 NA NA NA NA 15 20 14 NA ...
## ..$ 10     : num [1:56] NA 22 NA NA 19 25 24 28 23 NA ...
## ..$ 11     : num [1:56] 22 19 NA NA 30 30 NA 19 NA 17 ...
## ..$ 12     : num [1:56] 20 26 NA NA 33 18 NA NA 19 23 ...
## ..$ 13     : num [1:56] 23 NA NA NA 24 21 NA NA 26 NA ...
## ..$ 14     : num [1:56] 20 NA NA NA 44 24 NA 31 NA NA ...
## ..$ 15     : num [1:56] 26 17 NA 26 22 22 NA 37 NA NA ...
## ..$ 16     : num [1:56] 29 22 NA NA 20 21 20 NA NA NA ...
## ..$ 17     : num [1:56] 25 28 NA NA NA 18 15 NA NA NA ...
## $ PassingAttempts  :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 30 48 42 36 36 NA NA 39 NA ...
## ..$ 2      : num [1:56] 34 28 36 50 35 39 NA 32 24 NA ...
## ..$ 3      : num [1:56] NA 35 48 42 39 35 NA 47 23 NA ...

```

```

## ..$ 4      : num [1:56] 49 42 51 26 30 45 NA 49 NA NA ...
## ..$ 5      : num [1:56] 26 NA 44 29 NA 40 NA 36 30 NA ...
## ..$ 6      : num [1:56] NA 35 22 NA 25 38 NA 19 47 NA ...
## ..$ 7      : num [1:56] 39 33 NA NA 24 29 NA 39 27 21 ...
## ..$ 8      : num [1:56] 15 29 NA NA 31 47 NA NA 33 NA ...
## ..$ 9      : num [1:56] 52 38 NA NA NA NA 30 36 20 NA ...
## ..$ 10     : num [1:56] NA 29 NA NA 31 34 47 37 39 NA ...
## ..$ 11     : num [1:56] 28 27 NA NA 45 37 NA 35 NA 28 ...
## ..$ 12     : num [1:56] 32 35 NA NA 45 28 NA NA 36 34 ...
## ..$ 13     : num [1:56] 36 NA NA NA 40 30 NA NA 36 NA ...
## ..$ 14     : num [1:56] 35 NA NA NA 66 43 NA 46 NA NA ...
## ..$ 15     : num [1:56] 42 31 NA 45 30 35 NA 57 NA NA ...
## ..$ 16     : num [1:56] 38 36 NA NA 29 28 34 NA NA NA ...
## ..$ 17     : num [1:56] 47 45 NA NA NA 37 34 NA NA NA ...
## $ PassingCompletionPercentage:Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1       : num [1:56] NA 70 56.2 66.7 66.7 44.4 NA NA 66.7 NA ...
## ..$ 2       : num [1:56] 73.5 67.9 52.8 66 65.7 76.9 NA 68.8 66.7 NA ...
## ..$ 3       : num [1:56] NA 68.6 60.4 66.7 56.4 71.4 NA 74.5 78.3 NA ...
## ..$ 4       : num [1:56] 63.3 57.1 64.7 69.2 60 71.1 NA 61.2 NA NA ...
## ..$ 5       : num [1:56] 73.1 NA 63.6 65.5 NA 75 NA 58.3 76.7 NA ...
## ..$ 6       : num [1:56] NA 68.6 81.8 NA 68 52.6 NA 57.9 66 NA ...
## ..$ 7       : num [1:56] 69.2 69.7 NA NA 58.3 72.4 NA 48.7 63 61.9 ...
## ..$ 8       : num [1:56] 66.7 62.1 NA NA 54.8 68.1 NA NA 78.8 NA ...
## ..$ 9       : num [1:56] 65.4 63.2 NA NA NA NA 50 55.6 70 NA ...
## ..$ 10      : num [1:56] NA 75.9 NA NA 61.3 73.5 51.1 75.7 59 NA ...
## ..$ 11      : num [1:56] 78.6 70.4 NA NA 66.7 81.1 NA 54.3 NA 60.7 ...
## ..$ 12      : num [1:56] 62.5 74.3 NA NA 73.3 64.3 NA NA 52.8 67.6 ...
## ..$ 13      : num [1:56] 63.9 NA NA NA 60 70 NA NA 72.2 NA ...
## ..$ 14      : num [1:56] 57.1 NA NA NA 66.7 55.8 NA 67.4 NA NA ...
## ..$ 15      : num [1:56] 61.9 54.8 NA 57.8 73.3 62.9 NA 64.9 NA NA ...
## ..$ 16      : num [1:56] 76.3 61.1 NA NA 69 75 58.8 NA NA NA ...
## ..$ 17      : num [1:56] 53.2 62.2 NA NA NA 48.6 44.1 NA NA NA ...
## $ PassingYards :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1       : num [1:56] NA 321 268 311 263 267 NA NA 187 NA ...
## ..$ 2       : num [1:56] 217 252 332 343 243 447 NA 239 166 NA ...
## ..$ 3       : num [1:56] NA 294 325 313 235 378 NA 366 249 NA ...
## ..$ 4       : num [1:56] 235 242 357 179 216 307 NA 288 NA NA ...
## ..$ 5       : num [1:56] 222 NA 291 221 NA 303 NA 225 194 NA ...
## ..$ 6       : num [1:56] NA 248 283 NA 252 257 NA 128 354 NA ...
## ..$ 7       : num [1:56] 186 233 NA NA 224 249 NA 134 209 188 ...
## ..$ 8       : num [1:56] 101 254 NA NA 317 333 NA NA 257 NA ...
## ..$ 9       : num [1:56] 261 313 NA NA NA NA 201 220 140 NA ...
## ..$ 10      : num [1:56] NA 215 NA NA 236 266 273 273 262 NA ...
## ..$ 11      : num [1:56] 183 195 NA NA 299 340 NA 205 NA 282 ...
## ..$ 12      : num [1:56] 141 317 NA NA 351 227 NA NA 307 215 ...
## ..$ 13      : num [1:56] 269 NA NA NA 290 258 NA NA 331 NA ...
## ..$ 14      : num [1:56] 269 NA NA NA 506 233 NA 228 NA NA ...
## ..$ 15      : num [1:56] 288 212 NA 290 281 298 NA 434 NA NA ...
## ..$ 16      : num [1:56] 237 288 NA NA 226 224 209 NA NA NA ...
## ..$ 17      : num [1:56] 203 317 NA NA NA 190 145 NA NA NA ...
## $ PassingYardsPerAttempt :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...

```

```

## ..$ 1      : num [1:56] NA 10.7 5.6 7.4 7.3 7.4 NA NA 4.8 NA ...
## ..$ 2      : num [1:56] 6.4 9 9.2 6.9 6.9 11.5 NA 7.5 6.9 NA ...
## ..$ 3      : num [1:56] NA 8.4 6.8 7.5 6 10.8 NA 7.8 10.8 NA ...
## ..$ 4      : num [1:56] 4.8 5.8 7 6.9 7.2 6.8 NA 5.9 NA NA ...
## ..$ 5      : num [1:56] 8.5 NA 6.6 7.6 NA 7.6 NA 6.2 6.5 NA ...
## ..$ 6      : num [1:56] NA 7.1 12.9 NA 10.1 6.8 NA 6.7 7.5 NA ...
## ..$ 7      : num [1:56] 4.8 7.1 NA NA 9.3 8.6 NA 3.4 7.7 9 ...
## ..$ 8      : num [1:56] 6.7 8.8 NA NA 10.2 7.1 NA NA 7.8 NA ...
## ..$ 9      : num [1:56] 5 8.2 NA NA NA NA 6.7 6.1 7 NA ...
## ..$ 10     : num [1:56] NA 7.4 NA NA 7.6 7.8 5.8 7.4 6.7 NA ...
## ..$ 11     : num [1:56] 6.5 7.2 NA NA 6.6 9.2 NA 5.9 NA 10.1 ...
## ..$ 12     : num [1:56] 4.4 9.1 NA NA 7.8 8.1 NA NA 8.5 6.3 ...
## ..$ 13     : num [1:56] 7.5 NA NA NA 7.2 8.6 NA NA 9.2 NA ...
## ..$ 14     : num [1:56] 7.7 NA NA NA 7.7 5.4 NA 5 NA NA ...
## ..$ 15     : num [1:56] 6.9 6.8 NA 6.4 9.4 8.5 NA 7.6 NA NA ...
## ..$ 16     : num [1:56] 6.2 8 NA NA 7.8 8 6.1 NA NA NA ...
## ..$ 17     : num [1:56] 4.3 7 NA NA NA 5.1 4.3 NA NA NA ...
## $ PassingTouchdowns      :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 1 1 1 2 0 NA NA 0 NA ...
## ..$ 2      : num [1:56] 2 1 1 2 2 3 NA 1 2 NA ...
## ..$ 3      : num [1:56] NA 2 2 3 1 5 NA 3 1 NA ...
## ..$ 4      : num [1:56] 1 1 1 4 1 2 NA 2 NA NA ...
## ..$ 5      : num [1:56] 0 NA 1 3 NA 1 NA 2 2 NA ...
## ..$ 6      : num [1:56] NA 1 3 NA 1 2 NA 1 2 NA ...
## ..$ 7      : num [1:56] 1 1 NA NA 2 2 NA 1 3 2 ...
## ..$ 8      : num [1:56] 1 2 NA NA 1 1 NA NA 2 NA ...
## ..$ 9      : num [1:56] 2 2 NA NA NA NA 2 2 1 NA ...
## ..$ 10     : num [1:56] NA 2 NA NA 2 3 1 2 1 NA ...
## ..$ 11     : num [1:56] 1 2 NA NA 4 3 NA 0 NA 1 ...
## ..$ 12     : num [1:56] 0 1 NA NA 4 4 NA NA 3 1 ...
## ..$ 13     : num [1:56] 2 NA NA NA 2 0 NA NA 1 NA ...
## ..$ 14     : num [1:56] 2 NA NA NA 2 1 NA 1 NA NA ...
## ..$ 15     : num [1:56] 1 1 NA 3 2 1 NA 3 NA NA ...
## ..$ 16     : num [1:56] 2 1 NA NA 2 2 2 NA NA NA ...
## ..$ 17     : num [1:56] 2 1 NA NA NA 2 1 NA NA NA ...
## $ PassingInterceptions    :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 0 3 1 1 0 NA NA 2 NA ...
## ..$ 2      : num [1:56] 1 0 1 1 0 0 NA 1 0 NA ...
## ..$ 3      : num [1:56] NA 3 0 1 0 0 NA 2 0 NA ...
## ..$ 4      : num [1:56] 2 2 1 0 1 0 NA 0 NA NA ...
## ..$ 5      : num [1:56] 0 NA 0 0 NA 1 NA 1 1 NA ...
## ..$ 6      : num [1:56] NA 1 1 NA 1 1 NA 0 2 NA ...
## ..$ 7      : num [1:56] 0 0 NA NA 0 0 NA 0 1 1 ...
## ..$ 8      : num [1:56] 0 0 NA NA 1 0 NA NA 0 NA ...
## ..$ 9      : num [1:56] 2 1 NA NA NA NA 1 1 0 NA ...
## ..$ 10     : num [1:56] NA 1 NA NA 1 0 0 0 1 NA ...
## ..$ 11     : num [1:56] 1 0 NA NA 0 0 NA 0 NA 0 ...
## ..$ 12     : num [1:56] 0 0 NA NA 2 1 NA NA 0 2 ...
## ..$ 13     : num [1:56] 0 NA NA NA 1 1 NA NA 0 NA ...
## ..$ 14     : num [1:56] 1 NA NA NA 0 2 NA 2 NA NA ...
## ..$ 15     : num [1:56] 0 0 NA 3 1 1 NA 1 NA NA ...
## ..$ 16     : num [1:56] 0 1 NA NA 0 1 2 NA NA NA ...

```

```

## ..$ 17      : num [1:56] 1 0 NA NA NA 0 1 NA NA NA ...
## $ PassingRating      :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 116.1 53.1 86.5 95 ...
## ..$ 2      : num [1:56] 97.3 108 82.2 90.7 104.8 ...
## ..$ 3      : num [1:56] NA 77.6 94.5 102.6 82.8 ...
## ..$ 4      : num [1:56] 64.6 61.8 83.5 128 79.3 ...
## ..$ 5      : num [1:56] 98.6 NA 90.2 122.9 NA ...
## ..$ 6      : num [1:56] NA 86.4 139.4 NA 97.4 ...
## ..$ 7      : num [1:56] 88.2 99.7 NA NA 117.4 ...
## ..$ 8      : num [1:56] 107.9 113.3 NA NA 87.7 ...
## ..$ 9      : num [1:56] 74.3 95.6 NA NA NA ...
## ..$ 10     : num [1:56] NA 104.8 NA NA 92.9 ...
## ..$ 11     : num [1:56] 90.9 115.5 NA NA 115 ...
## ..$ 12     : num [1:56] 72.5 111.2 NA NA 106.8 ...
## ..$ 13     : num [1:56] 105 NA NA NA 88.5 ...
## ..$ 14     : num [1:56] 88.9 NA NA NA 99.7 ...
## ..$ 15     : num [1:56] 90.2 87 NA 71.5 110.6 ...
## ..$ 16     : num [1:56] 109 84 NA NA 115 ...
## ..$ 17     : num [1:56] 69.7 90.7 NA NA NA ...
## $ RushingAttempts    :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 3 1 7 3 2 NA NA 2 NA ...
## ..$ 2      : num [1:56] 1 1 3 2 2 2 NA 0 4 NA ...
## ..$ 3      : num [1:56] NA 1 3 4 1 1 NA 0 4 NA ...
## ..$ 4      : num [1:56] 0 1 2 1 1 1 NA 3 NA NA ...
## ..$ 5      : num [1:56] 3 NA 0 4 NA 2 NA 0 3 NA ...
## ..$ 6      : num [1:56] NA 0 5 NA 2 1 NA 1 3 NA ...
## ..$ 7      : num [1:56] 1 3 NA NA 3 5 NA 0 3 0 ...
## ..$ 8      : num [1:56] 1 6 NA NA 3 1 NA NA 1 NA ...
## ..$ 9      : num [1:56] 1 1 NA NA NA NA 4 0 5 NA ...
## ..$ 10     : num [1:56] NA 0 NA NA 2 1 1 0 0 NA ...
## ..$ 11     : num [1:56] 2 3 NA NA 1 0 NA 2 NA 1 ...
## ..$ 12     : num [1:56] 6 1 NA NA 3 5 NA NA 2 2 ...
## ..$ 13     : num [1:56] 1 NA NA NA 1 0 NA NA 7 NA ...
## ..$ 14     : num [1:56] 0 NA NA NA 1 0 NA 0 NA NA ...
## ..$ 15     : num [1:56] 4 3 NA 6 3 2 NA 0 NA NA ...
## ..$ 16     : num [1:56] 4 2 NA NA 1 2 2 NA NA NA ...
## ..$ 17     : num [1:56] 0 3 NA NA NA 0 1 NA NA NA ...
## $ RushingYards       :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 11 2 21 -8 0 NA NA 0 NA ...
## ..$ 2      : num [1:56] -1 8 6 8 -2 9 NA 0 31 NA ...
## ..$ 3      : num [1:56] NA -1 7 23 6 6 NA 0 7 NA ...
## ..$ 4      : num [1:56] 0 8 1 -1 -1 2 NA 22 NA NA ...
## ..$ 5      : num [1:56] -4 NA 0 32 NA 5 NA 0 -3 NA ...
## ..$ 6      : num [1:56] NA 0 -4 NA -2 -1 NA 0 21 NA ...
## ..$ 7      : num [1:56] 0 37 NA NA -3 5 NA 0 1 0 ...
## ..$ 8      : num [1:56] 9 17 NA NA -2 2 NA NA 0 NA ...
## ..$ 9      : num [1:56] 7 -1 NA NA NA NA 2 0 13 NA ...
## ..$ 10     : num [1:56] NA 0 NA NA 5 0 1 0 0 NA ...
## ..$ 11     : num [1:56] -1 12 NA NA 10 0 NA 5 NA 0 ...
## ..$ 12     : num [1:56] 42 1 NA NA 25 -4 NA NA 28 9 ...
## ..$ 13     : num [1:56] -1 NA NA NA 4 0 NA NA 19 NA ...

```

```

## ..$ 14      : num [1:56] 0 NA NA NA 5 0 NA 0 NA NA ...
## ..$ 15      : num [1:56] 3 29 NA 43 10 -2 NA 0 NA NA ...
## ..$ 16      : num [1:56] -2 7 NA NA -1 6 1 NA NA NA ...
## ..$ 17      : num [1:56] 0 14 NA NA NA 0 -1 NA NA NA ...
## $ RushingYardsPerAttempt      :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1        : num [1:56] NA 3.7 2 3 -2.7 0 NA NA 0 NA ...
## ..$ 2        : num [1:56] -1 8 2 4 -1 4.5 NA 0 7.8 NA ...
## ..$ 3        : num [1:56] NA -1 2.3 5.8 6 6 NA 0 1.8 NA ...
## ..$ 4        : num [1:56] 0 8 0.5 -1 -1 2 NA 7.3 NA NA ...
## ..$ 5        : num [1:56] -1.3 NA 0 8 NA 2.5 NA 0 -1 NA ...
## ..$ 6        : num [1:56] NA 0 -0.8 NA -1 -1 NA 0 7 NA ...
## ..$ 7        : num [1:56] 0 12.3 NA NA -1 1 NA 0 0.3 0 ...
## ..$ 8        : num [1:56] 9 2.8 NA NA -0.7 2 NA NA 0 NA ...
## ..$ 9        : num [1:56] 7 -1 NA NA NA NA 0.5 0 2.6 NA ...
## ..$ 10       : num [1:56] NA 0 NA NA 2.5 0 1 0 0 NA ...
## ..$ 11       : num [1:56] -0.5 4 NA NA 10 0 NA 2.5 NA 0 ...
## ..$ 12       : num [1:56] 7 1 NA NA 8.3 -0.8 NA NA 14 4.5 ...
## ..$ 13       : num [1:56] -1 NA NA NA 4 0 NA NA 2.7 NA ...
## ..$ 14       : num [1:56] 0 NA NA NA 5 0 NA 0 NA NA ...
## ..$ 15       : num [1:56] 0.8 9.7 NA 7.2 3.3 -1 NA 0 NA NA ...
## ..$ 16       : num [1:56] -0.5 3.5 NA NA -1 3 0.5 NA NA NA ...
## ..$ 17       : num [1:56] 0 4.7 NA NA NA 0 -1 NA NA NA ...
## $ RushingTouchdowns          :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1        : num [1:56] NA 0 0 0 0 0 NA NA 1 NA ...
## ..$ 2        : num [1:56] 0 0 0 0 0 0 NA 0 0 NA ...
## ..$ 3        : num [1:56] NA 0 0 0 0 0 NA 0 0 NA ...
## ..$ 4        : num [1:56] 0 0 0 0 0 0 NA 1 NA NA ...
## ..$ 5        : num [1:56] 0 NA 0 0 NA 0 NA 0 0 NA ...
## ..$ 6        : num [1:56] NA 0 0 NA 0 0 NA 0 0 NA ...
## ..$ 7        : num [1:56] 0 0 NA NA 0 0 NA 0 1 0 ...
## ..$ 8        : num [1:56] 0 0 NA NA 0 0 NA NA 0 NA ...
## ..$ 9        : num [1:56] 0 0 NA NA NA NA 0 0 1 NA ...
## ..$ 10       : num [1:56] NA 0 NA NA 0 0 0 0 0 NA ...
## ..$ 11       : num [1:56] 0 0 NA NA 0 0 NA 0 NA 0 ...
## ..$ 12       : num [1:56] 0 0 NA NA 0 0 NA NA 0 0 ...
## ..$ 13       : num [1:56] 0 NA NA NA 0 0 NA NA 2 NA ...
## ..$ 14       : num [1:56] 0 NA NA NA 0 0 NA 0 NA NA ...
## ..$ 15       : num [1:56] 1 0 NA 0 0 0 NA 0 NA NA ...
## ..$ 16       : num [1:56] 0 0 NA NA 0 0 0 NA NA NA ...
## ..$ 17       : num [1:56] 0 0 NA NA NA 0 0 NA NA NA ...
## $ FumblesLost                :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1        : num [1:56] NA 0 0 0 0 0 NA NA 0 NA ...
## ..$ 2        : num [1:56] 0 0 0 1 0 0 NA 0 1 NA ...
## ..$ 3        : num [1:56] NA 0 0 0 1 1 NA 0 0 NA ...
## ..$ 4        : num [1:56] 0 1 0 0 0 0 NA 0 NA NA ...
## ..$ 5        : num [1:56] 0 NA 0 0 NA 1 NA 1 0 NA ...
## ..$ 6        : num [1:56] NA 0 0 NA 0 0 NA 0 0 NA ...
## ..$ 7        : num [1:56] 0 0 NA NA 0 0 NA 1 0 0 ...
## ..$ 8        : num [1:56] 0 2 NA NA 0 0 NA NA 0 NA ...
## ..$ 9        : num [1:56] 0 0 NA NA NA NA 0 1 0 NA ...
## ..$ 10       : num [1:56] NA 0 NA NA 0 0 0 1 0 NA ...

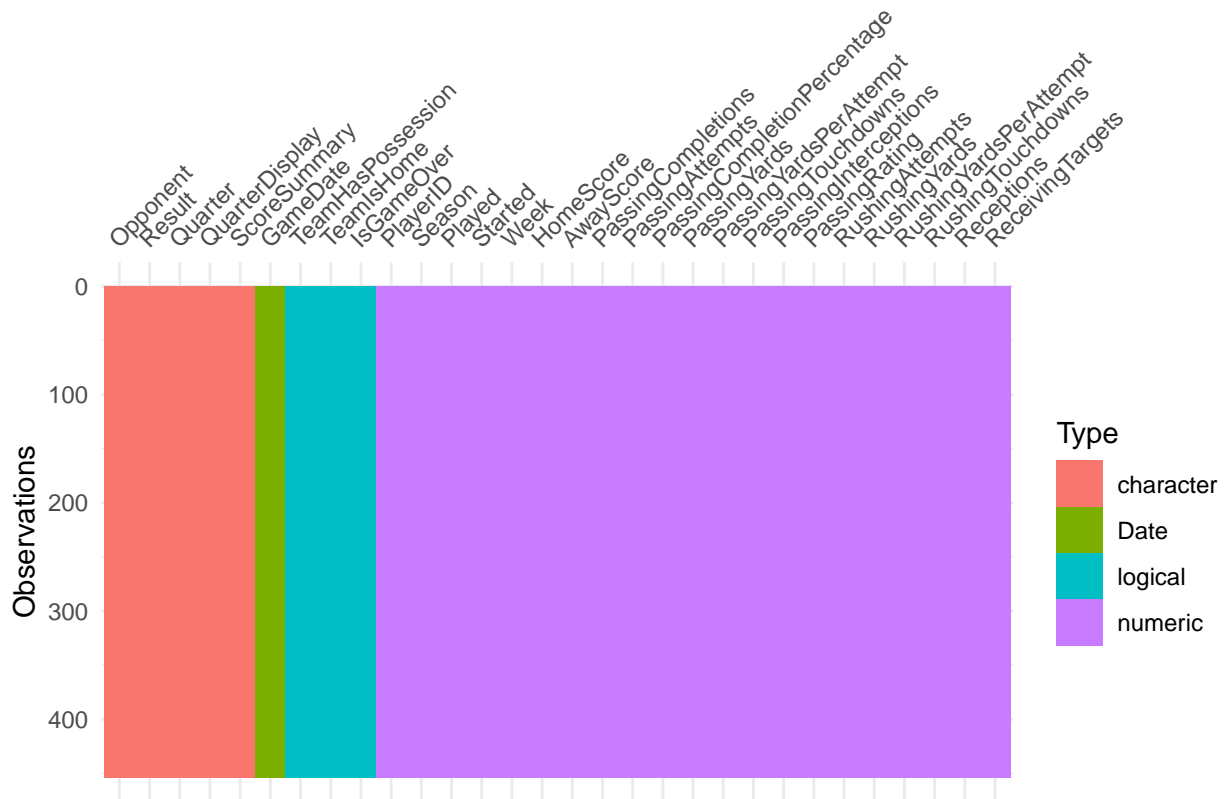
```

```
## ..$ 11      : num [1:56] 0 0 NA NA 0 0 NA 0 NA 0 ...
## ..$ 12      : num [1:56] 0 0 NA NA 0 1 NA NA 1 0 ...
## ..$ 13      : num [1:56] 0 NA NA NA 0 0 NA NA 0 NA ...
## ..$ 14      : num [1:56] 0 NA NA NA 0 0 NA 0 NA NA ...
## ..$ 15      : num [1:56] 0 0 NA 0 0 0 NA 0 NA NA ...
## ..$ 16      : num [1:56] 0 0 NA NA 0 0 0 NA NA NA ...
## ..$ 17      : num [1:56] 0 0 NA NA NA 0 0 NA NA NA ...
## $ FantasyPoints      :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : num [1:56] NA 17.94 8.92 16.54 15.72 ...
## ..$ 2      : num [1:56] 14.6 14.9 15.9 18.5 17.5 ...
## ..$ 3      : num [1:56] NA 13.7 21.7 24.8 12 ...
## ..$ 4      : num [1:56] 9.4 8.48 16.38 23.06 10.54 ...
## ..$ 5      : num [1:56] 8.48 NA 15.64 24.04 NA ...
## ..$ 6      : num [1:56] NA 11.9 20.9 NA 11.9 ...
## ..$ 7      : num [1:56] 11.4 17 NA NA 16.7 ...
## ..$ 8      : num [1:56] 8.94 15.86 NA NA 14.48 ...
## ..$ 9      : num [1:56] 15.1 18.4 NA NA NA ...
## ..$ 10     : num [1:56] NA 14.6 NA NA 17.9 ...
## ..$ 11     : num [1:56] 9.22 17 NA NA 28.96 ...
## ..$ 12     : num [1:56] 9.84 16.78 NA NA 30.54 ...
## ..$ 13     : num [1:56] 18.7 NA NA NA 18 ...
## ..$ 14     : num [1:56] 16.8 NA NA NA 28.7 ...
## ..$ 15     : num [1:56] 21.8 15.4 NA 21.9 18.2 ...
## ..$ 16     : num [1:56] 17.3 14.2 NA NA 16.9 ...
## ..$ 17     : num [1:56] 14.1 18.1 NA NA NA ...
## $ Team      :Classes 'tbl_df', 'tbl' and 'data.frame': 56 obs. of  18 variables:
## ..$ PlayerID: num [1:56] 611 732 2428 2593 3807 ...
## ..$ 1      : chr [1:56] NA "ATL" "ARI" "GB" ...
## ..$ 2      : chr [1:56] "BAL" "ATL" "ARI" "GB" ...
## ..$ 3      : chr [1:56] NA "ATL" "ARI" "GB" ...
## ..$ 4      : chr [1:56] "BAL" "ATL" "ARI" "GB" ...
## ..$ 5      : chr [1:56] "BAL" NA "ARI" "GB" ...
## ..$ 6      : chr [1:56] NA "ATL" "ARI" NA ...
## ..$ 7      : chr [1:56] "BAL" "ATL" NA NA ...
## ..$ 8      : chr [1:56] "BAL" "ATL" NA NA ...
## ..$ 9      : chr [1:56] "BAL" "ATL" NA NA ...
## ..$ 10     : chr [1:56] NA "ATL" NA NA ...
## ..$ 11     : chr [1:56] "BAL" "ATL" NA NA ...
## ..$ 12     : chr [1:56] "BAL" "ATL" NA NA ...
## ..$ 13     : chr [1:56] "BAL" NA NA NA ...
## ..$ 14     : chr [1:56] "BAL" NA NA NA ...
## ..$ 15     : chr [1:56] "BAL" "ATL" NA "GB" ...
## ..$ 16     : chr [1:56] "BAL" "ATL" NA NA ...
## ..$ 17     : chr [1:56] "BAL" "ATL" NA NA ...
```

## 1.3: Missing Data

### 1.3.1 Quarterbacks (Old Data QBdata dataset)

```
vis_dat(QBdata[1:30])
```

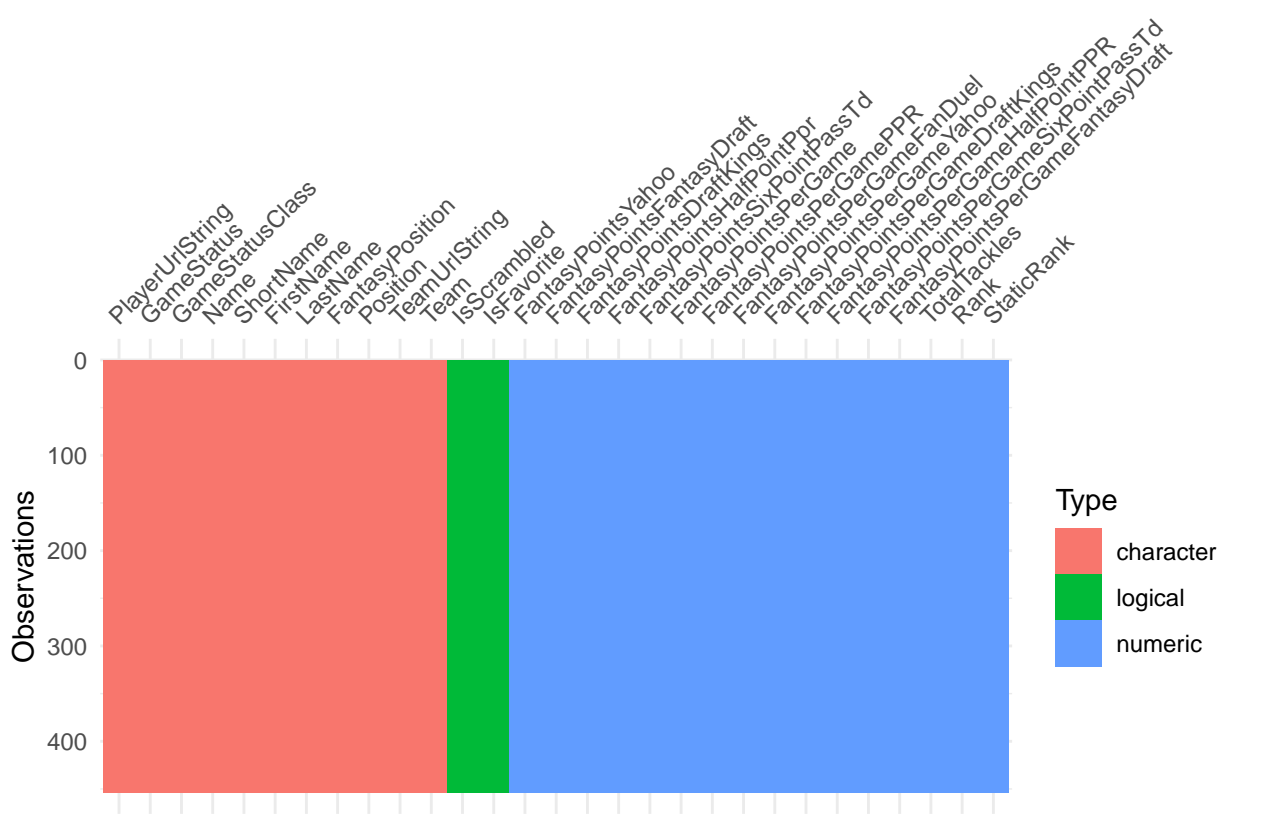


```
vis_dat(QBdata[31:60])
```



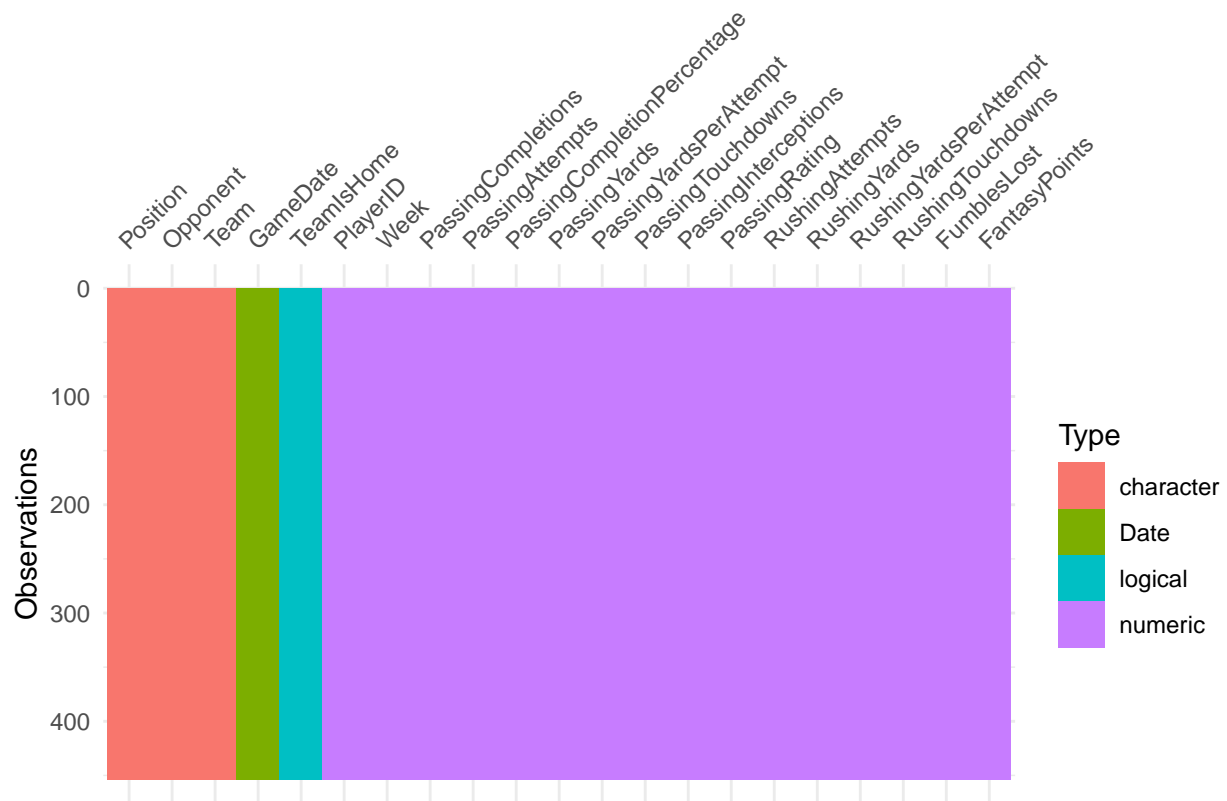


```
vis_dat(QBdata[61:89])
```



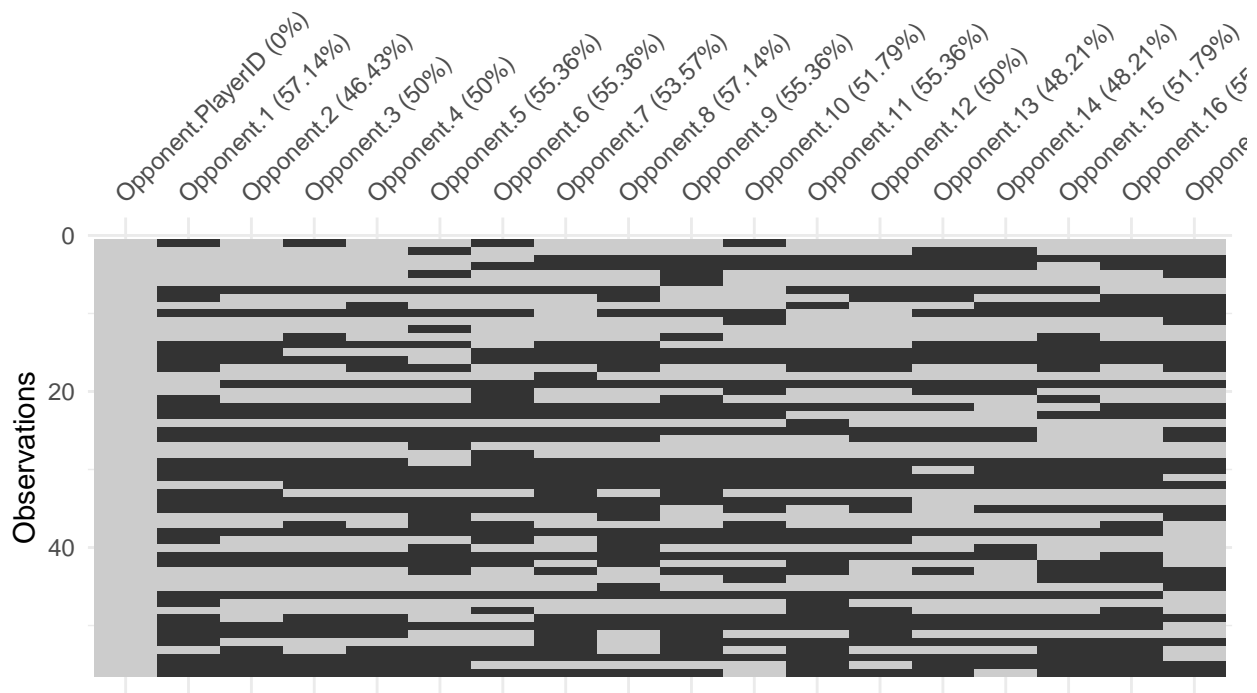
### 1.3.2 QBCrossSectional dataset (New dataset)

```
vis_dat(QBCrossSectional)
```

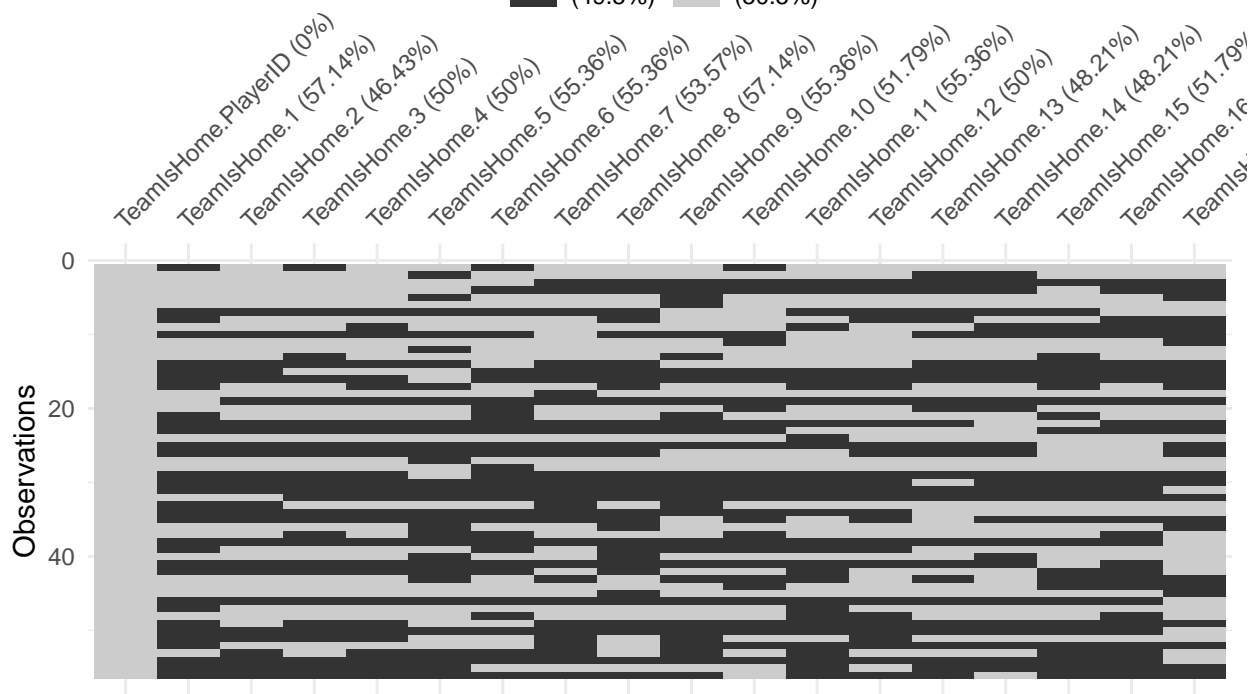


### 1.3.3 QBPanels dataset (New dataset)

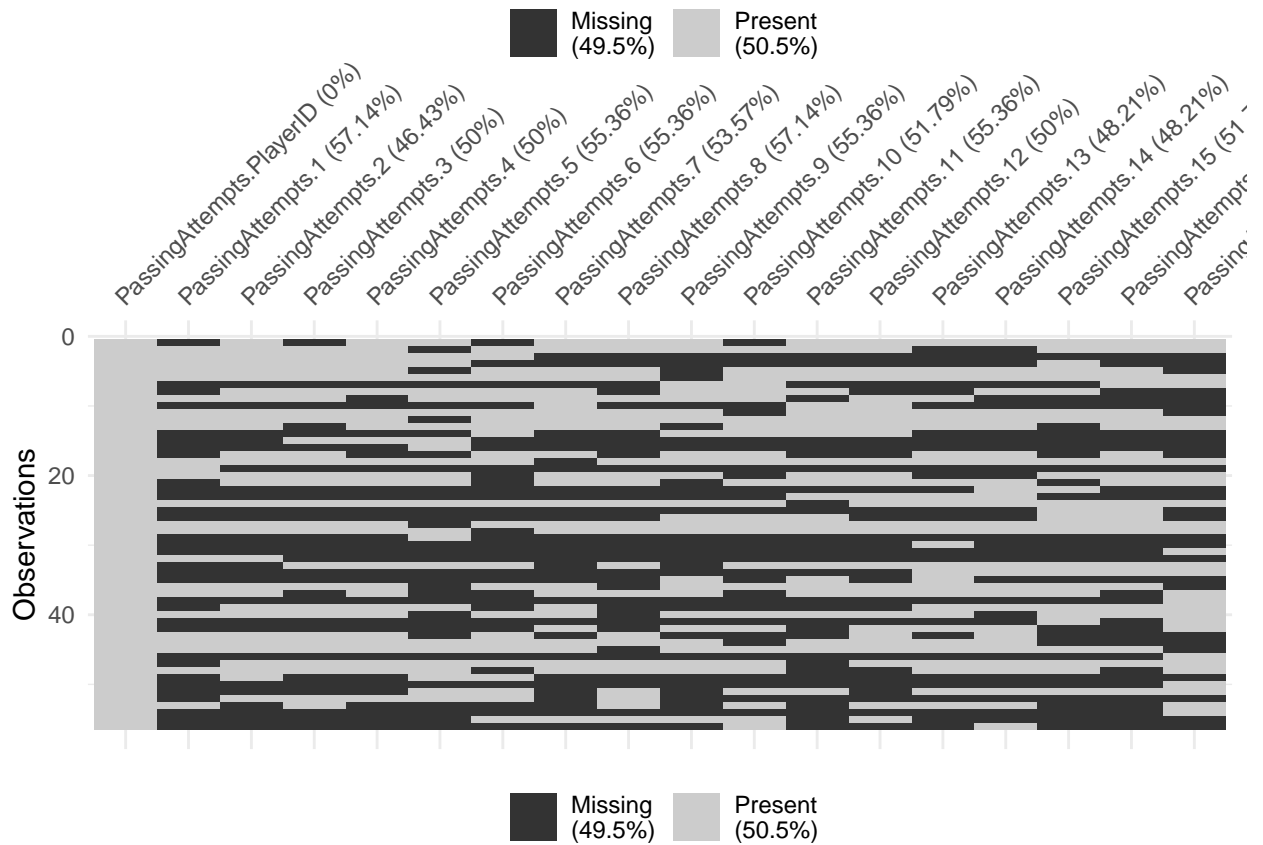
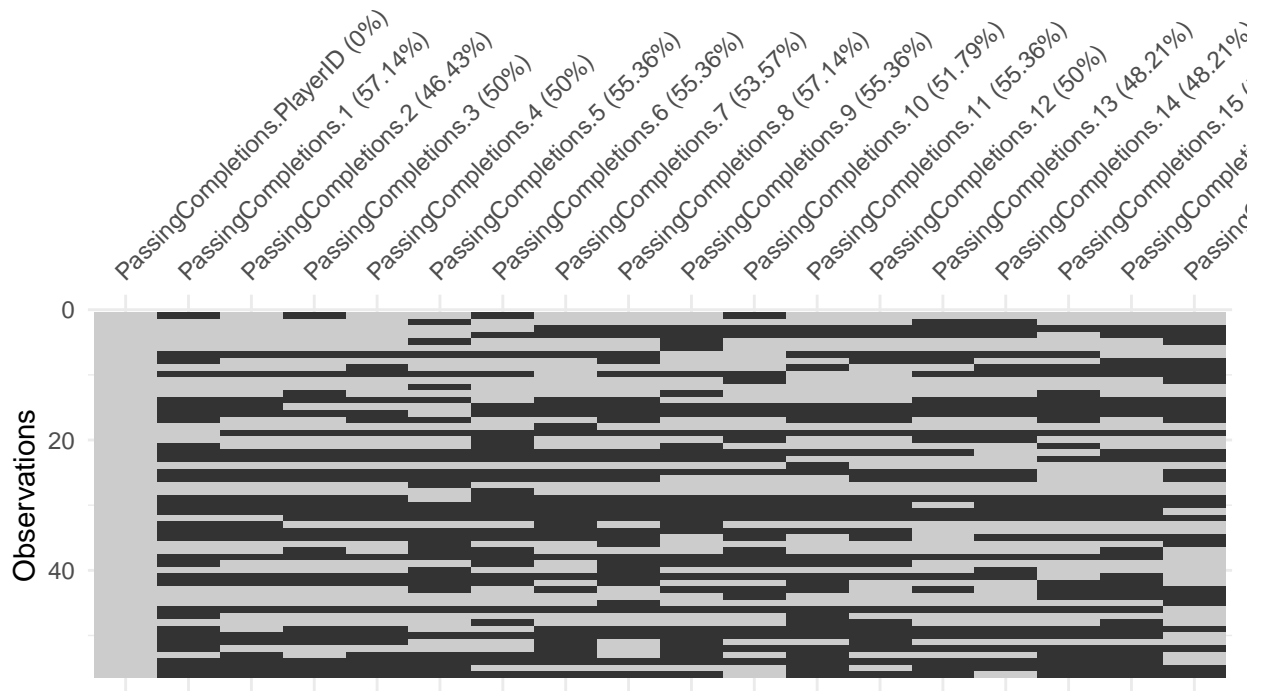
```
for (panel in 1:length(QBPanels)) {
  missing_plot <- vis_miss(as.data.frame(QBPanels[panel]))
  print(missing_plot)
}
```

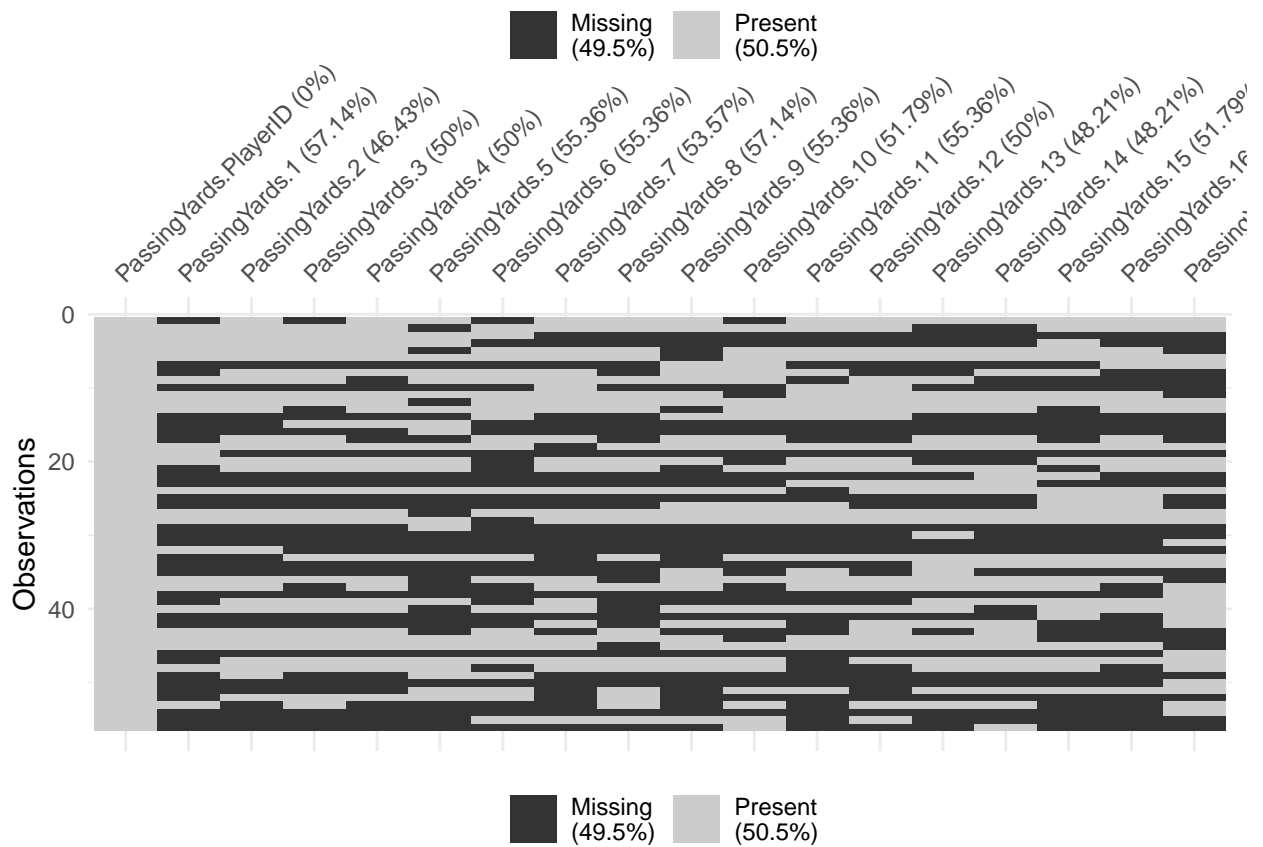
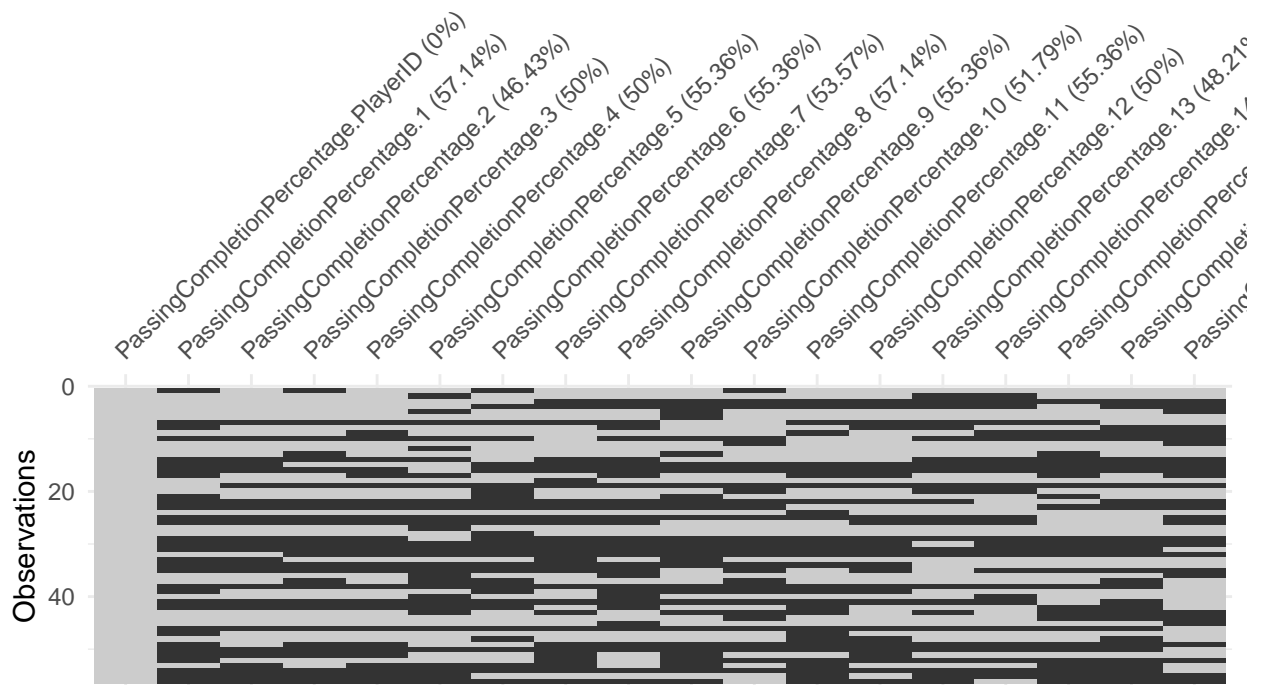


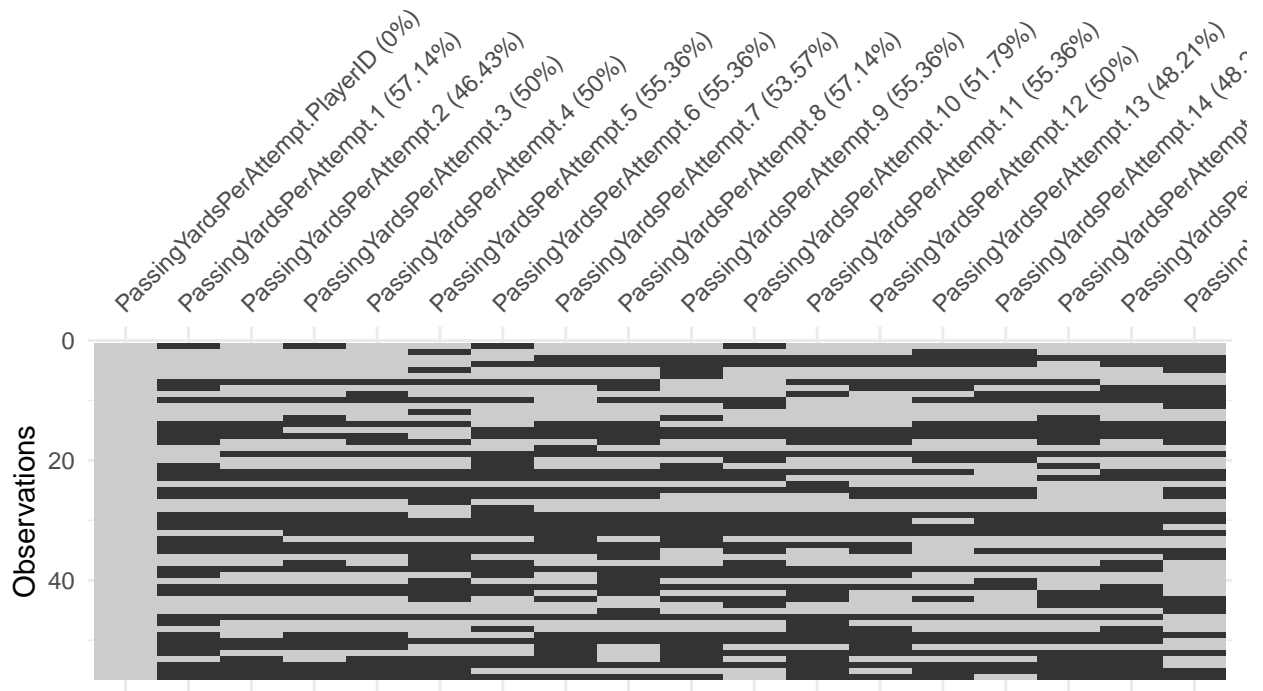
Missing (49.5%) Present (50.5%)



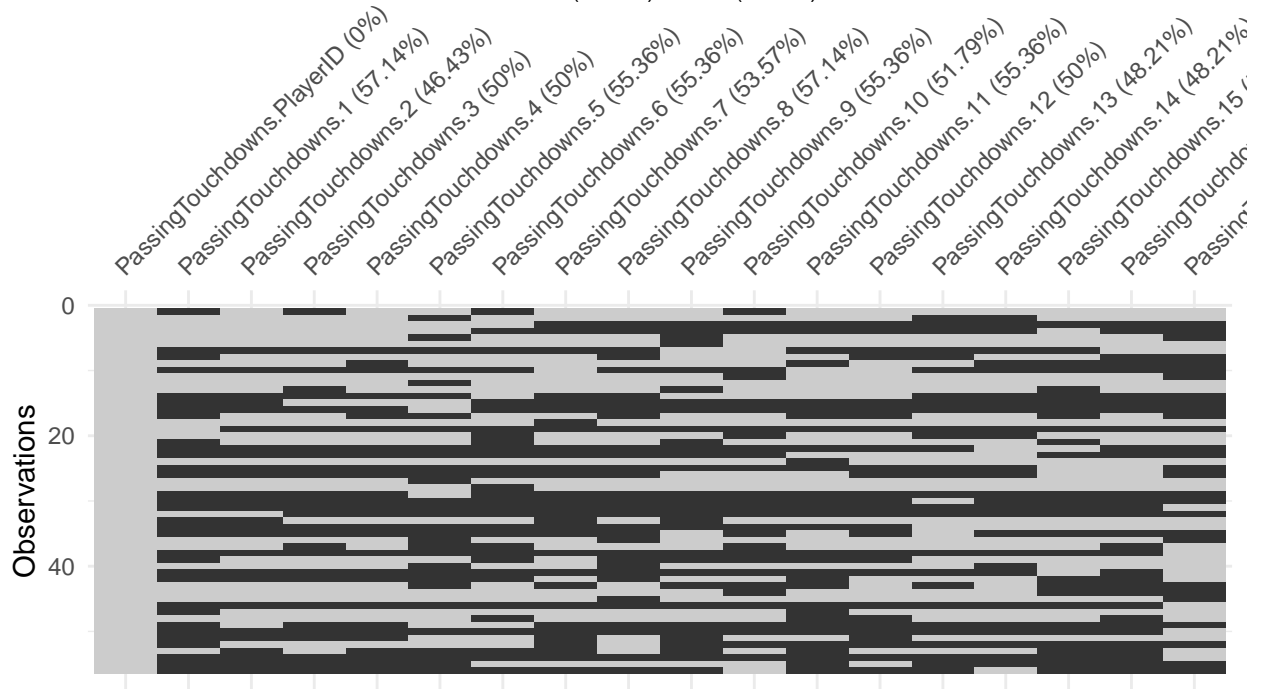
Missing (49.5%) Present (50.5%)



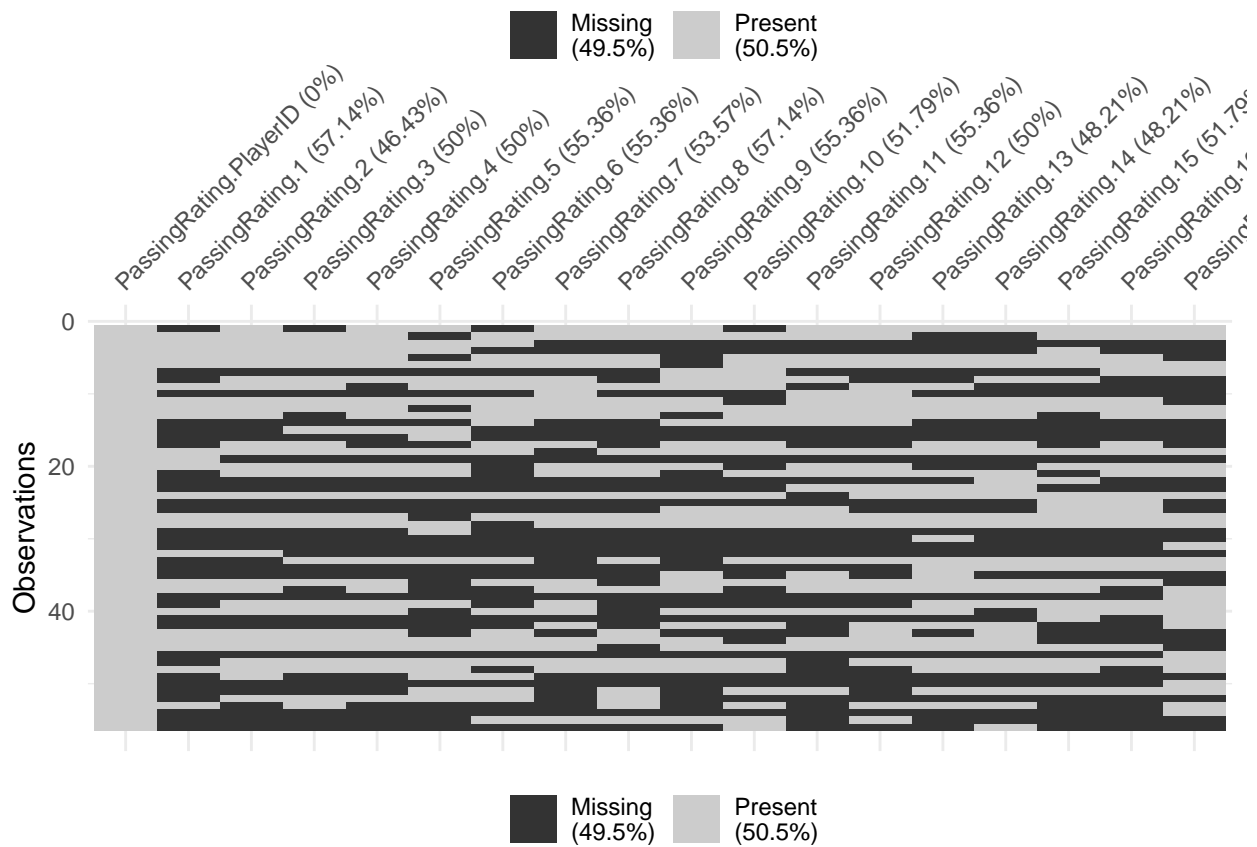
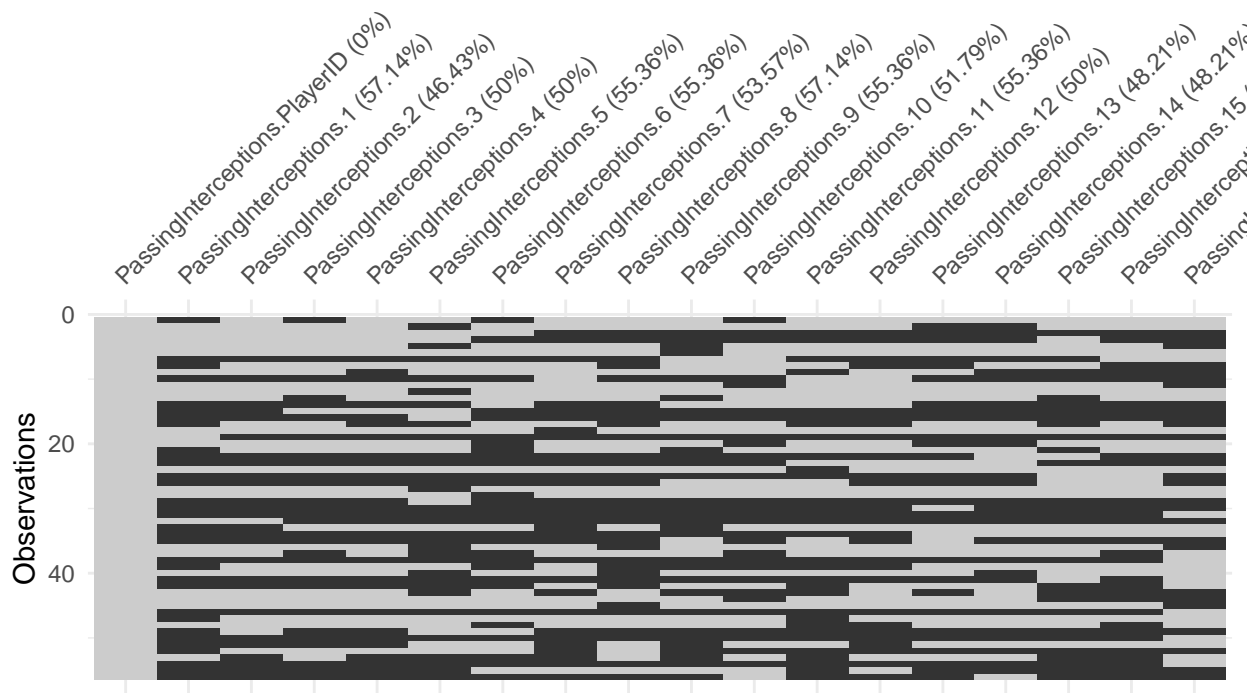


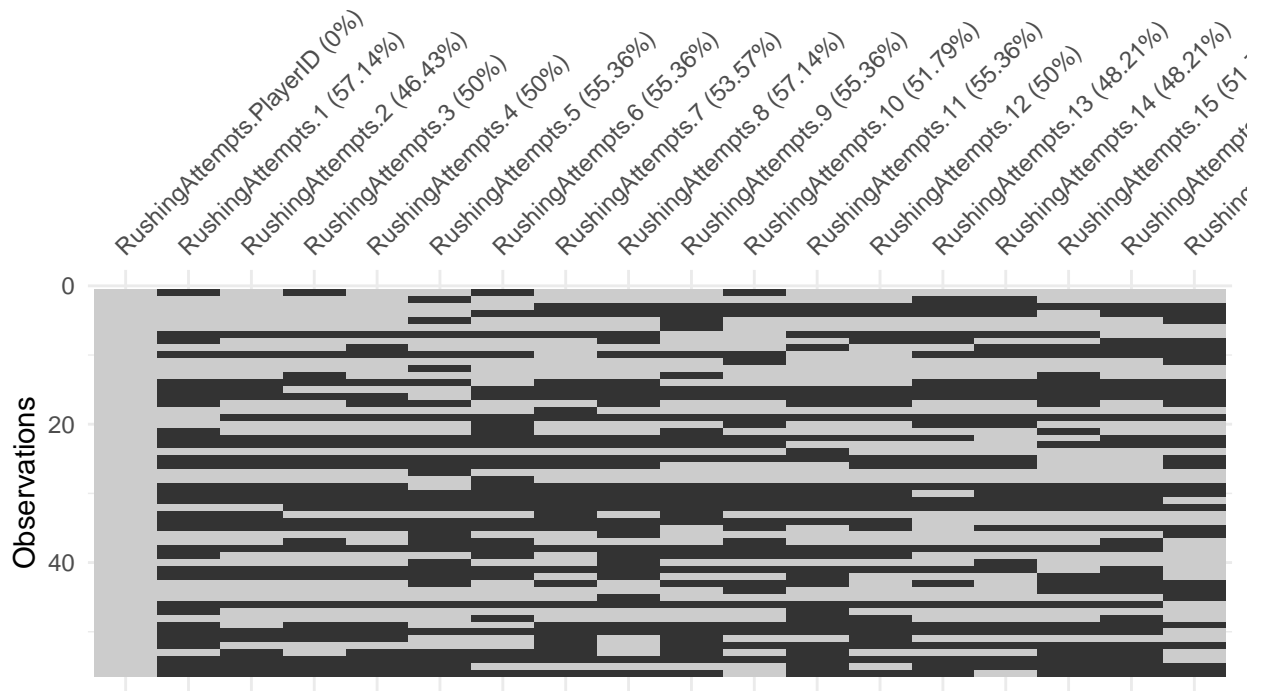


Missing (49.5%)
  Present (50.5%)

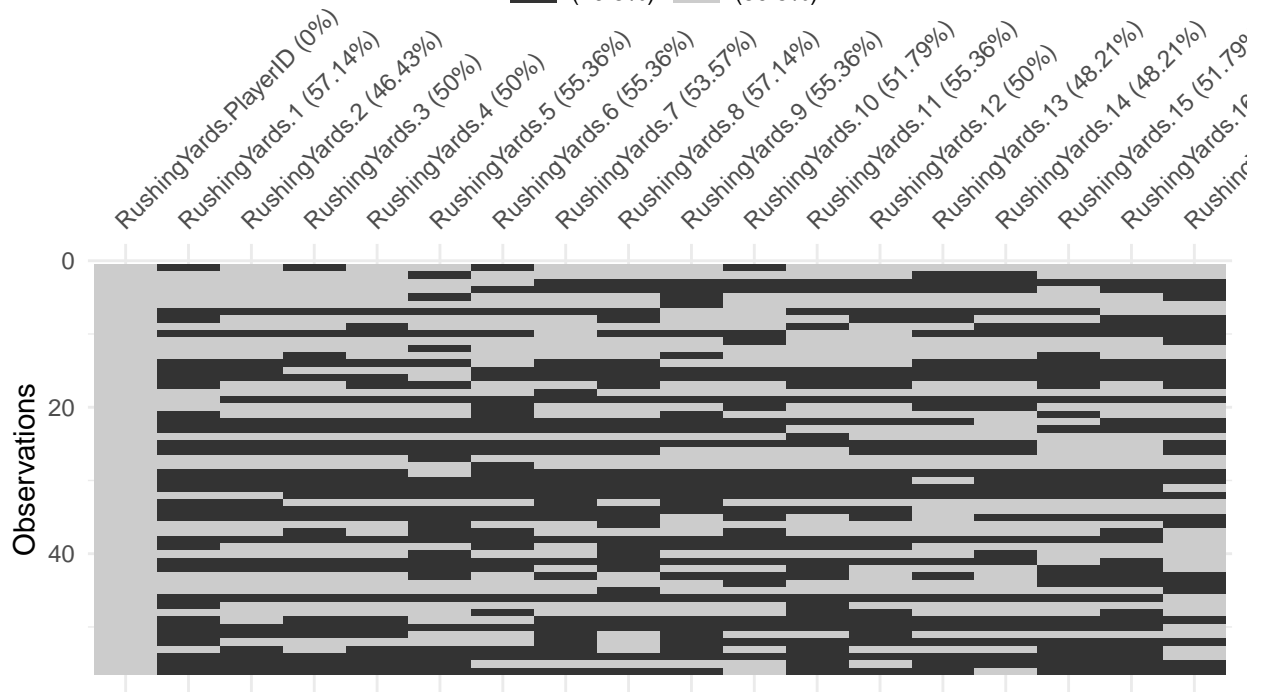


Missing (49.5%)
  Present (50.5%)



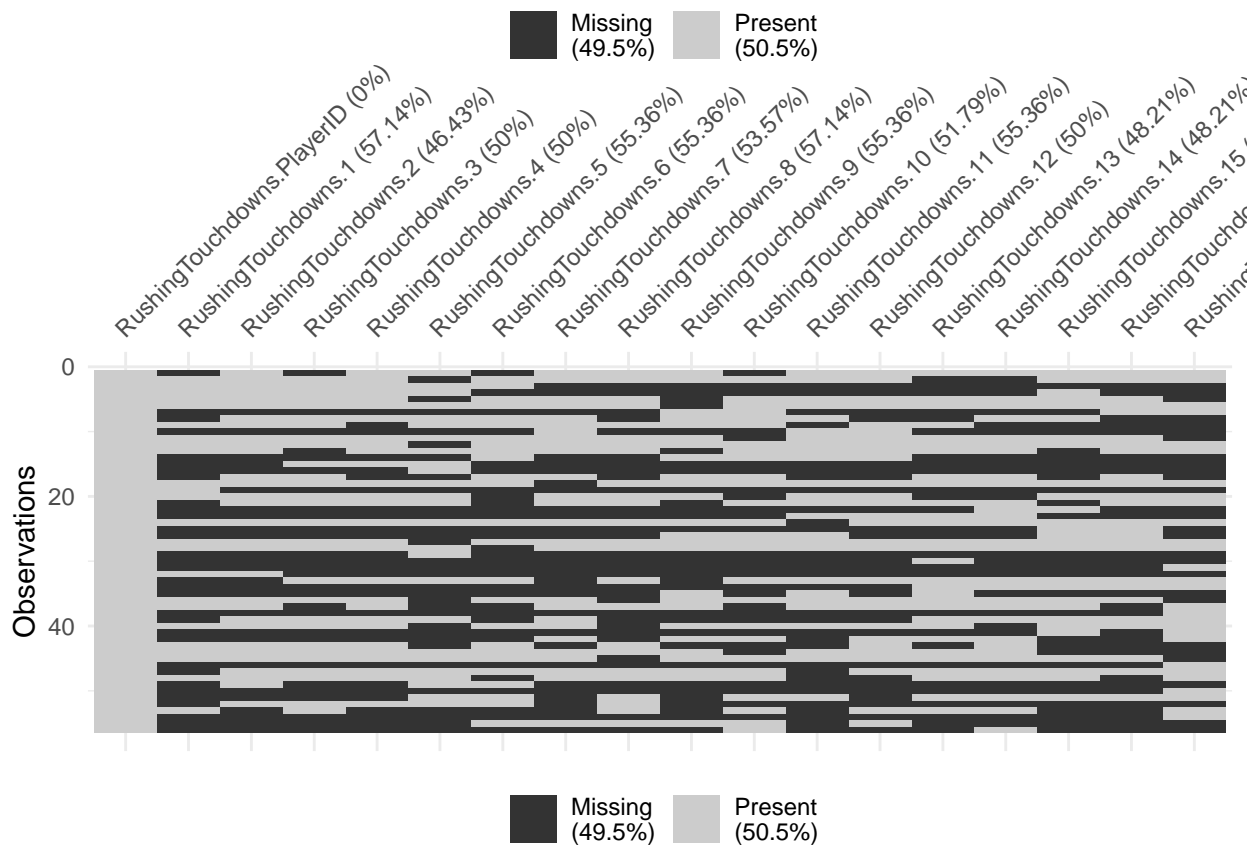
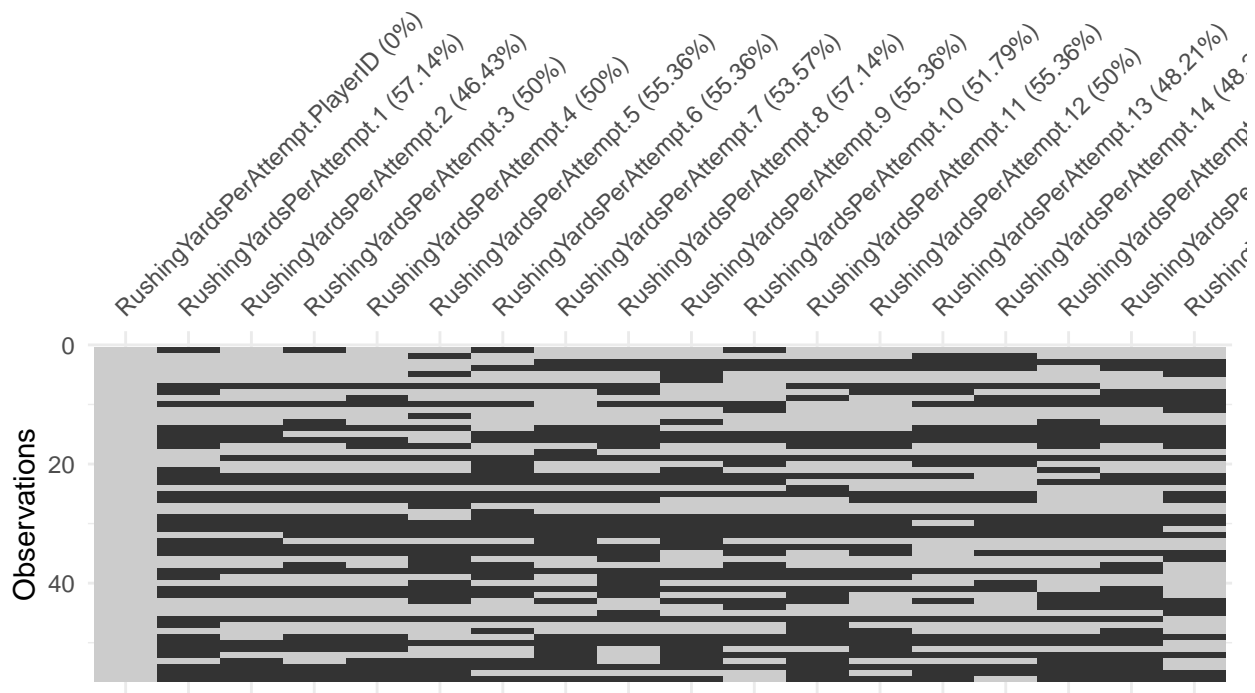


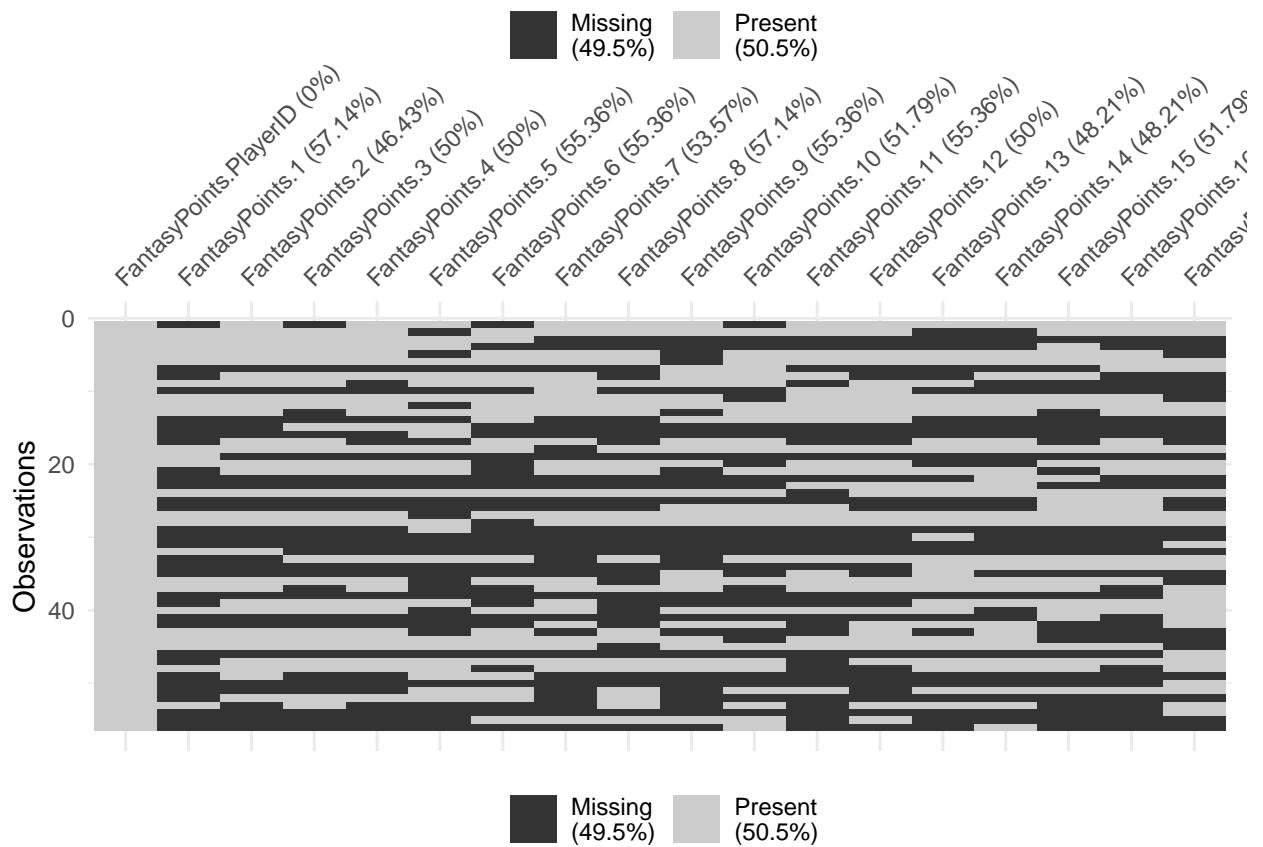
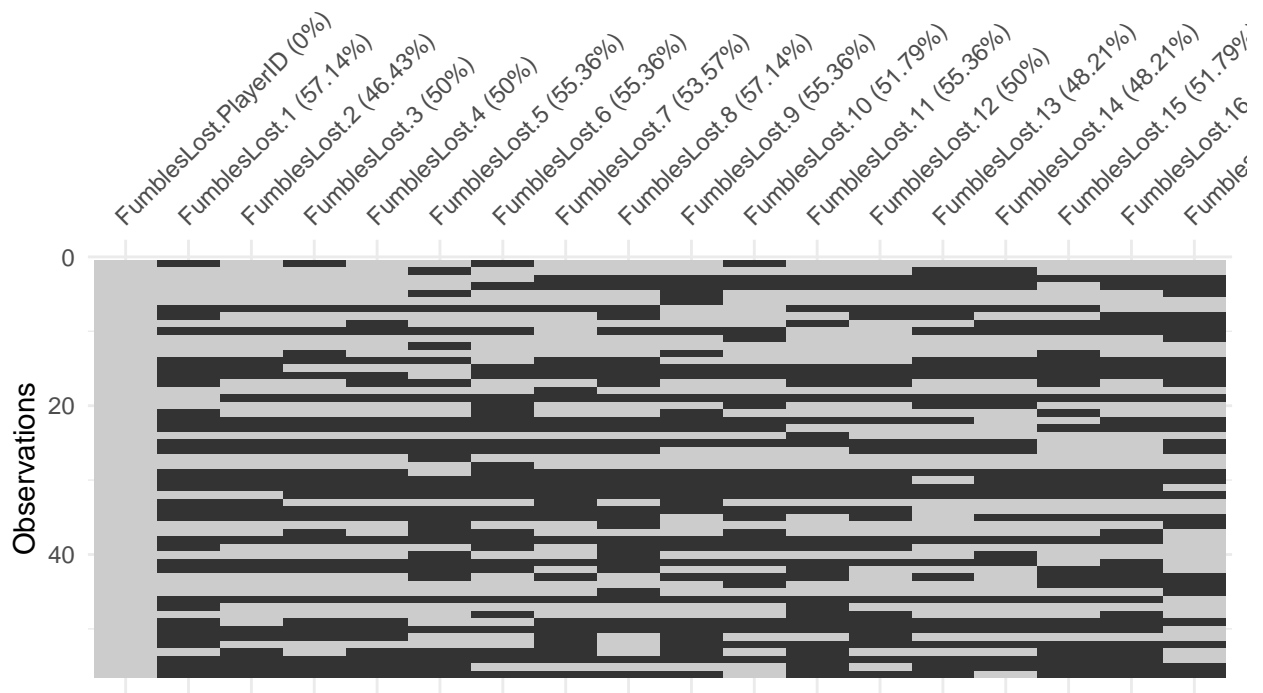
Missing (49.5%) Present (50.5%)

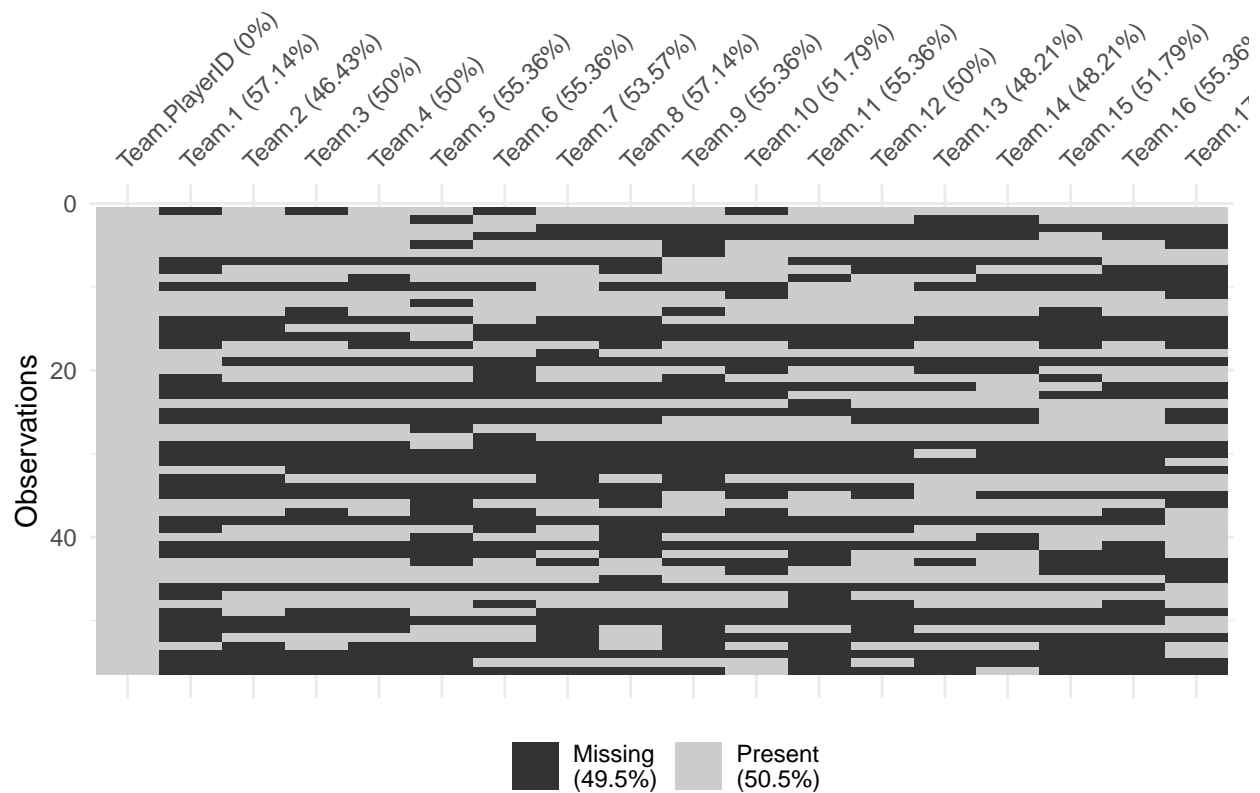


Missing (49.5%) Present (50.5%)









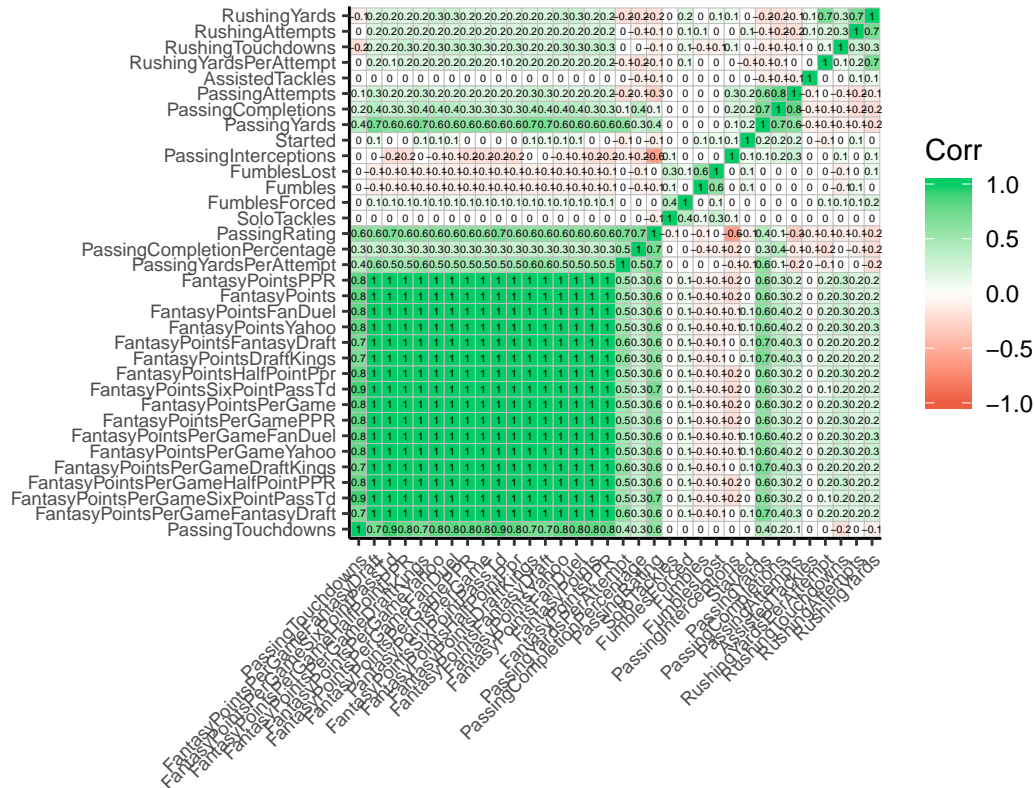
## 1.4: Corrrelogram

### 1.4.1 Corrrrelogram Style1 QBData (Old dataset)

```
QBdata_num <- QBdata[-c(1:3,5:16,29:36,39:52,55:56,74:89)]
corr <- round(cor(QBdata_num), 1)

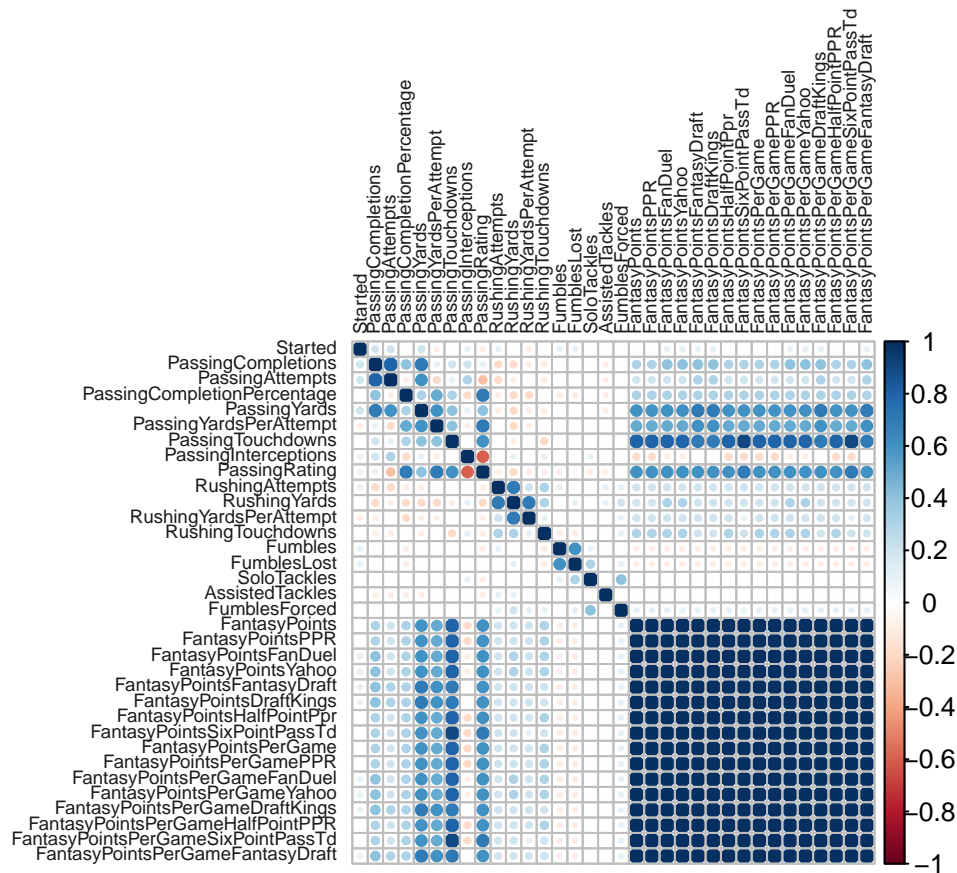
ggcorrplot(corr, hc.order = TRUE,
  type = "full",
  lab = TRUE,
  lab_size = 1.5,
  method="square",
  colors = c("tomato2", "white", "springgreen3"),
  title="Correlogram of Quarterbacks", tl.cex = 7,pch=2,pch.col =3,show.diag = T,
  ggtheme=theme_classic)
```

## Correlogram of Quarterbacks



### 1.4.2 Correlogram Style2 QBData (Old dataset)

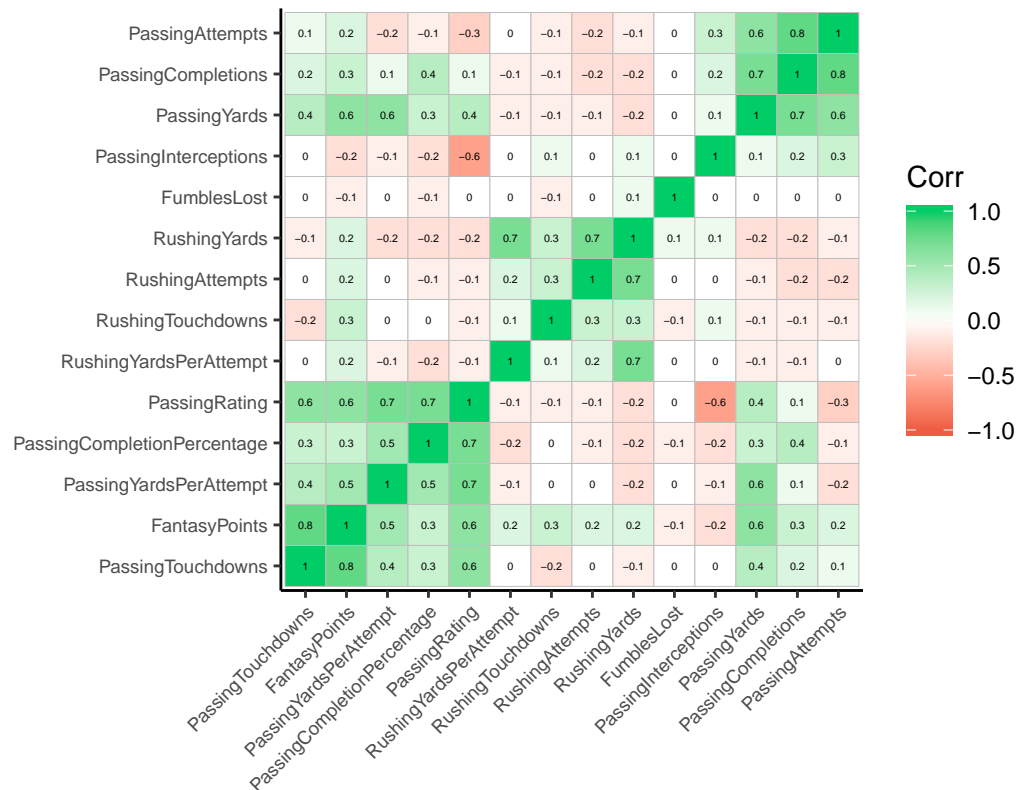
```
corrplot(corr, method="circle",tl.cex = 0.55,tl.col = "#1C1C1C")
```



### 1.4.3 Correlogram of Crosssectional data (New dataset)

```
QBX_ds <- QBCrossSectional[-c(1:6,21)]
corr <- round(cor(QBX_ds), 1)
ggcorrplot(corr, hc.order = TRUE,
  type = "full",
  lab = TRUE,
  lab_size = 1.5,
  method="square",
  colors = c("tomato2", "white", "springgreen3"),
  title="Correlogram of Quarterbacks", tl.cex = 7,pch=2,pch.col =3,show.diag = T,
  ggtheme=theme_classic)
```

## Correlogram of Quarterbacks

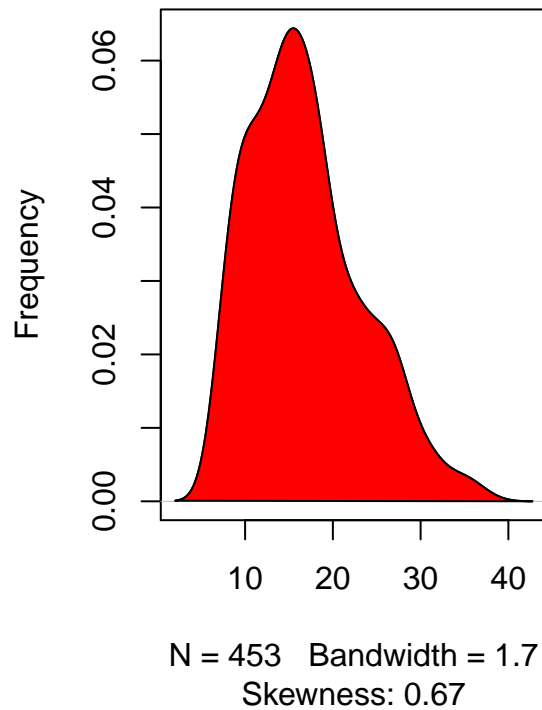


## 1.5: Distributions

### 1.5.1 Density plot for FantasyPoints is approximately Normal (Old QBData)

```
library(e1071)
par(mfrow=c(1, 2)) # divide graph area in 2 columns
target <- QBdata$FantasyPoints
plot(density(target), main="QBdata: FantasyPoints", ylab="Frequency", sub=paste("Skewness:", round(e1071::skewness(target), 2)))
polygon(density(target), col="red", asp=1.2 )
```

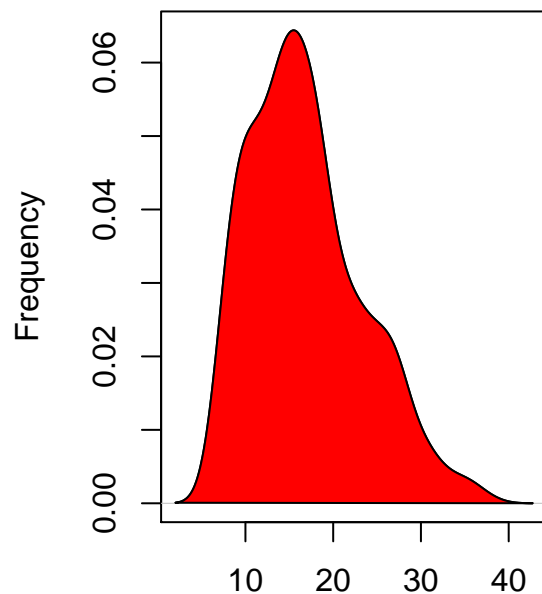
### QBdata: FantasyPoints



1.5.2 Density plot for Fantasypoints is approxmiately Normal (New QBCrossSectional)

```
par(mfrow=c(1, 2)) # divide graph area in 2 columns
target <- QBCrossSectional$FantasyPoints
plot(density(target), main="CrossSectional Dataset: FantasyPoints", ylab="Frequency", sub=paste("Skewness: 0.67"))
polygon(density(target), col="red")
```

## CrossSectional Dataset: FantasyPo

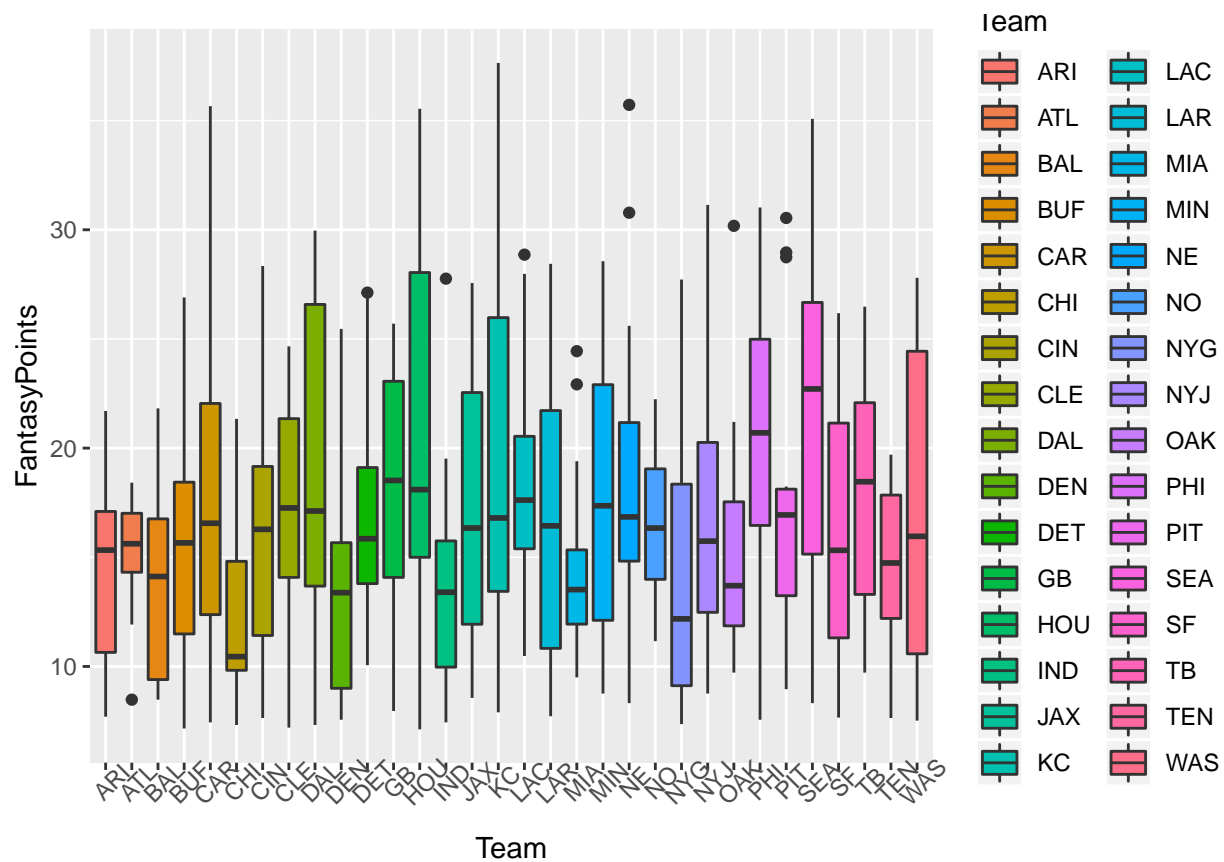


N = 453 Bandwidth = 1.7  
Skewness: 0.67

### 1.5.3 Boxplots - Target and Individual Predictor Behavior for per Team

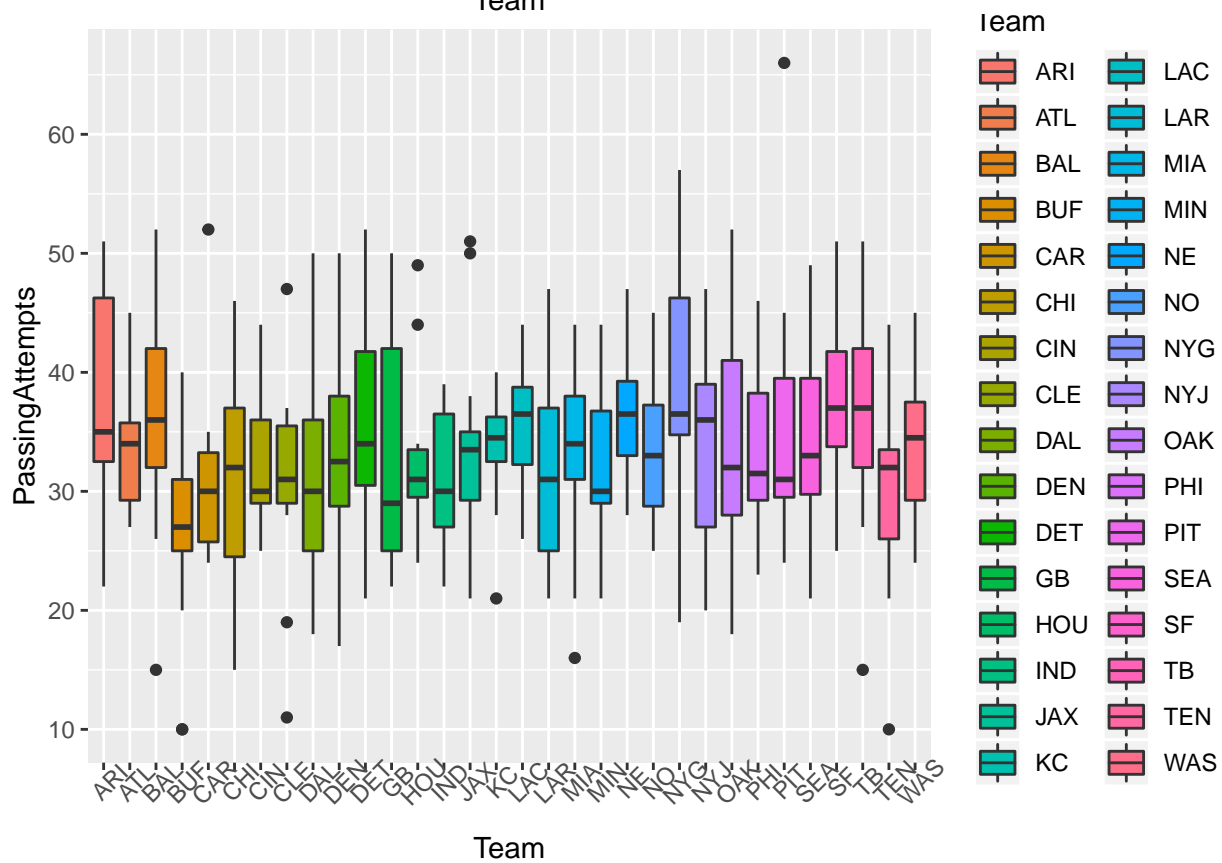
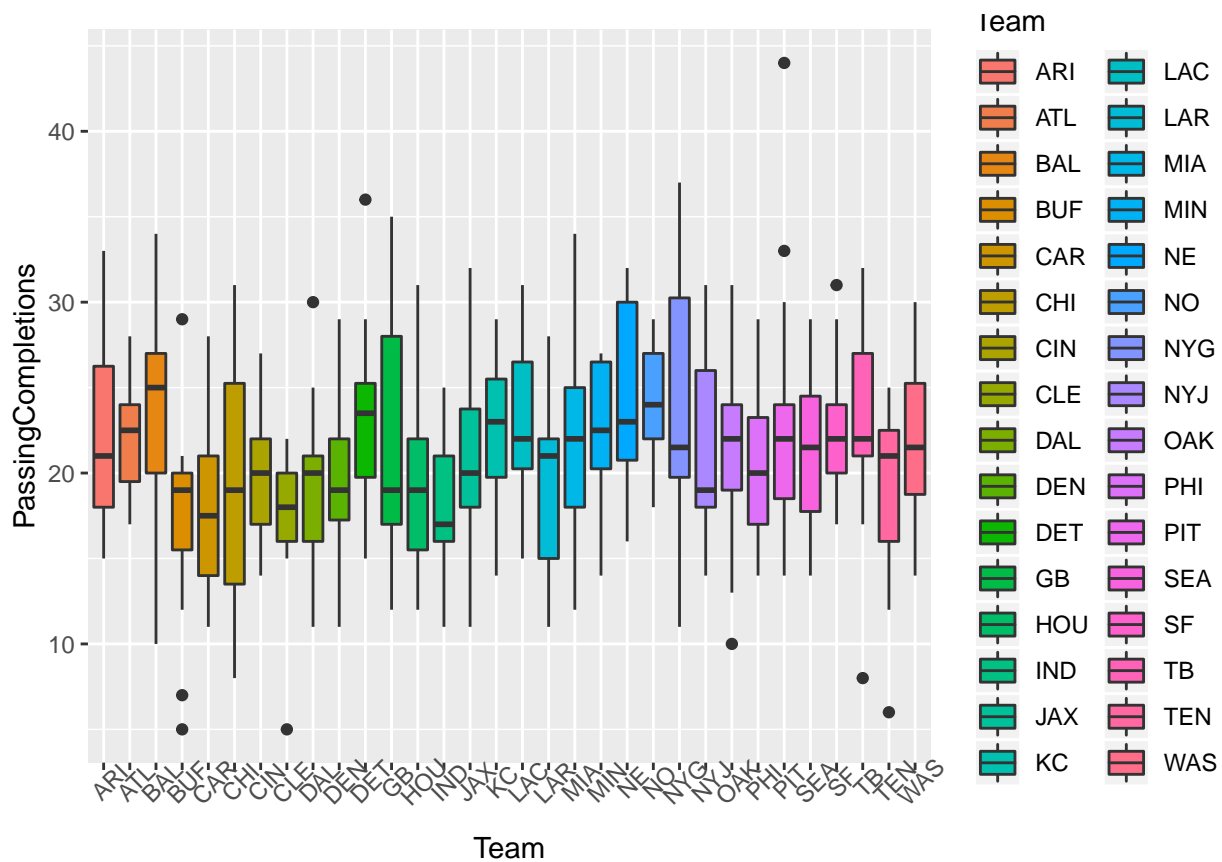
```
QBCrossSectional %>% ggplot(aes(y=FantasyPoints,x=Team,fill=Team))+  
  geom_boxplot()+  
  xlab("Team")+ylab("FantasyPoints")+  
  theme(axis.text.x = element_text(angle=45))
```

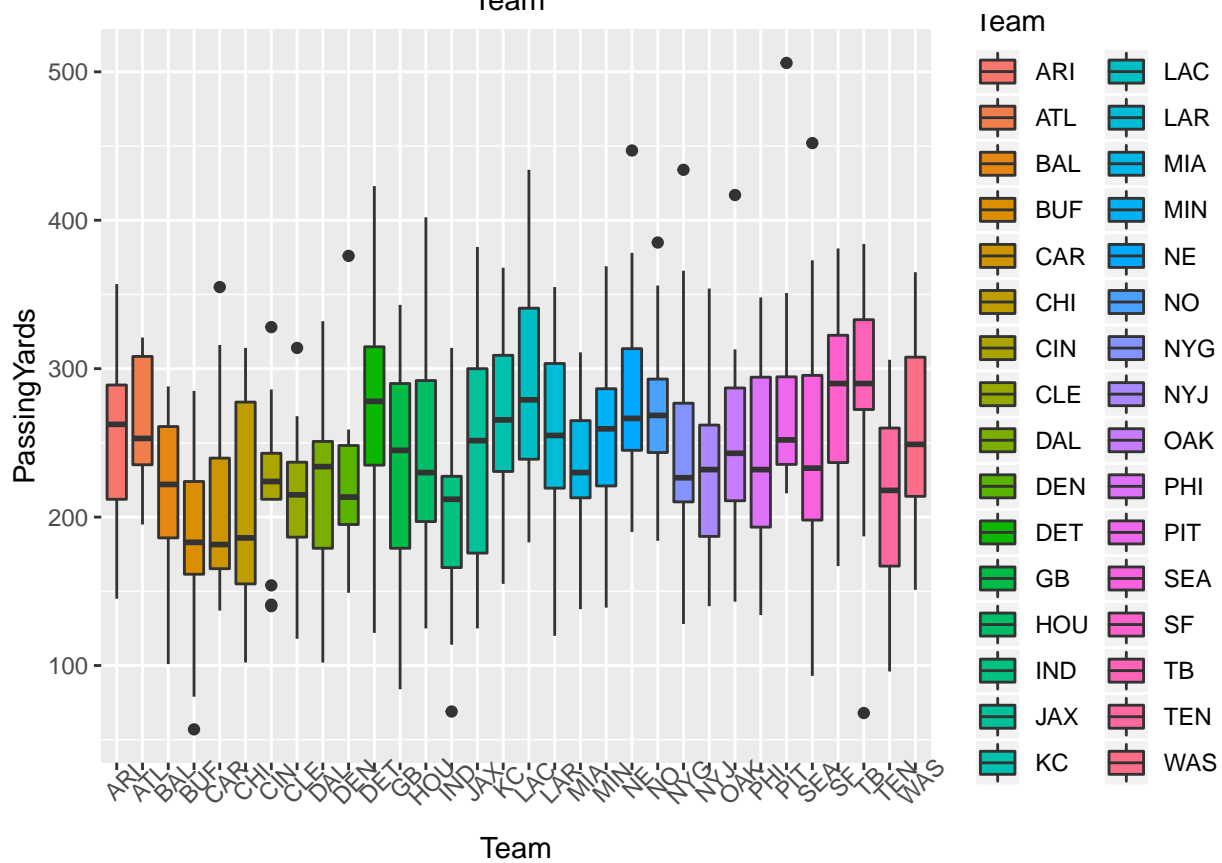
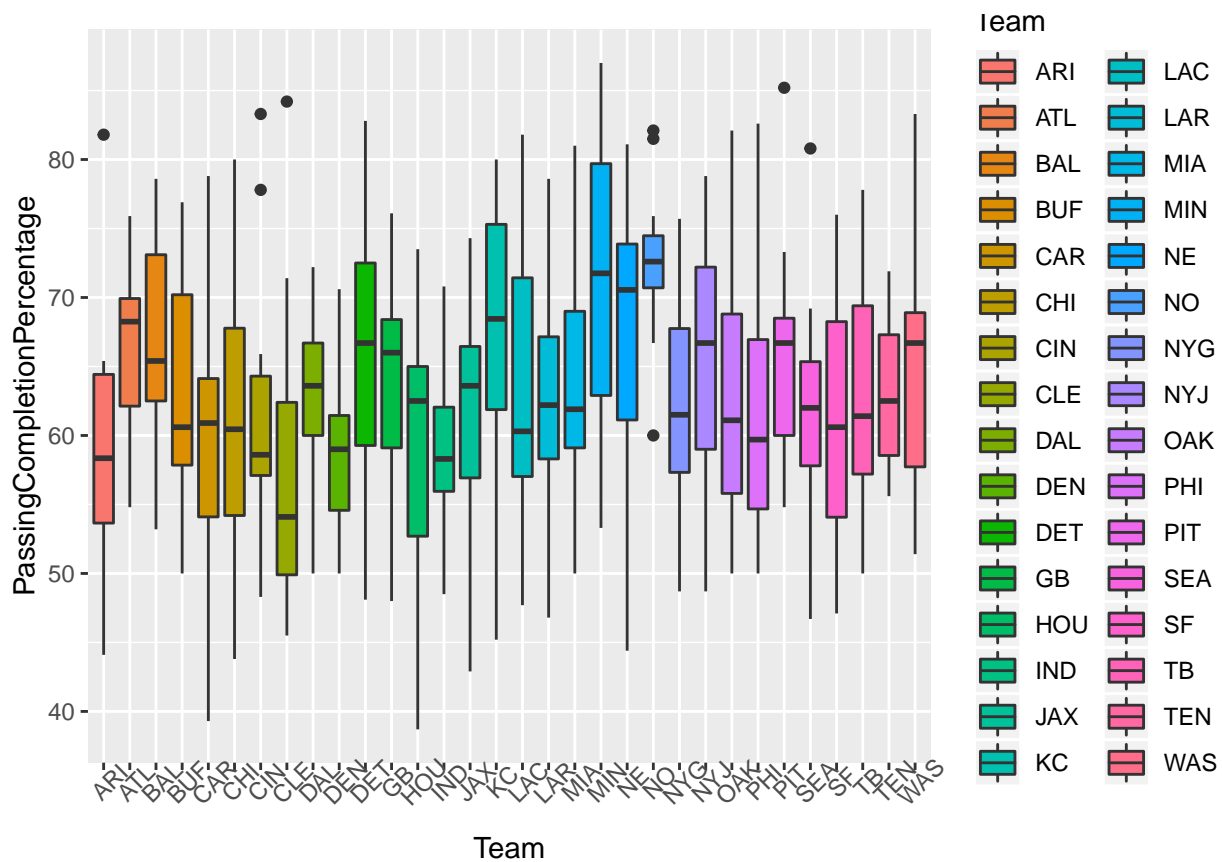


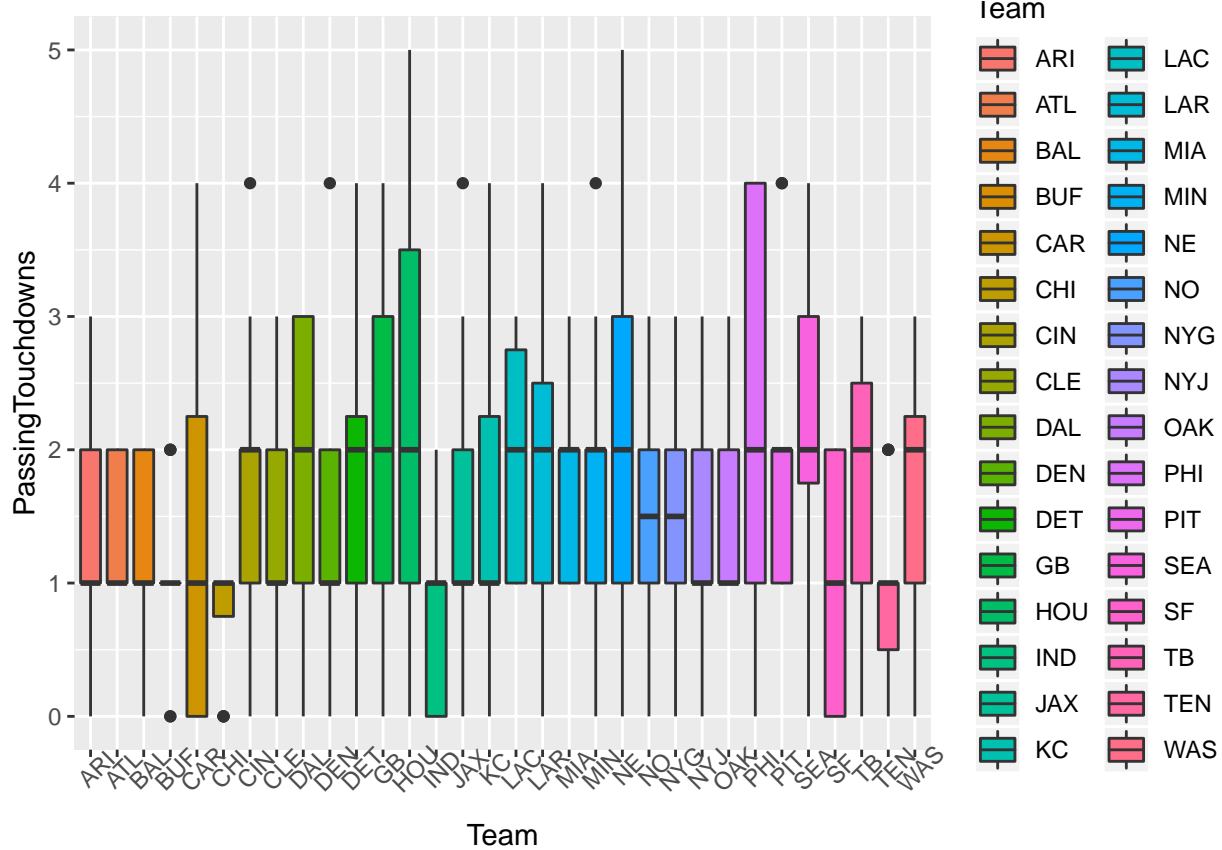
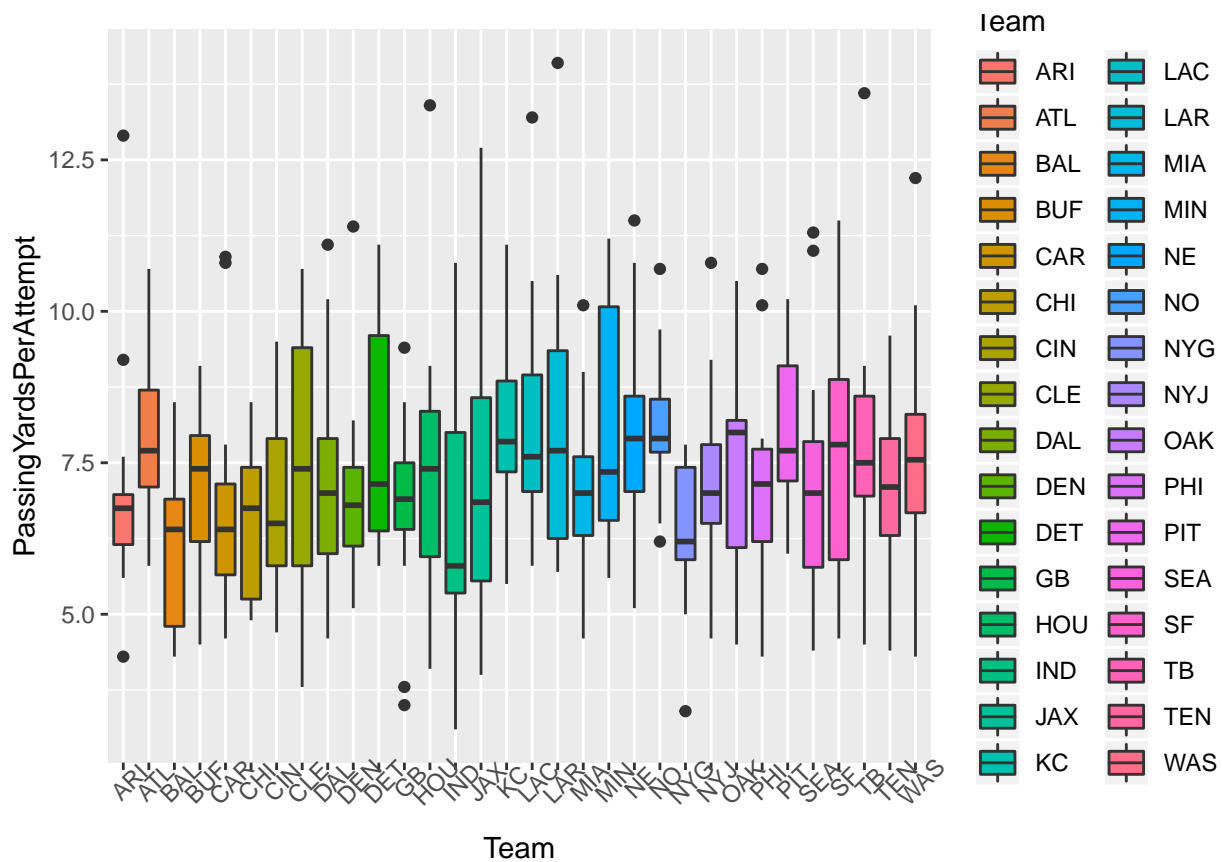


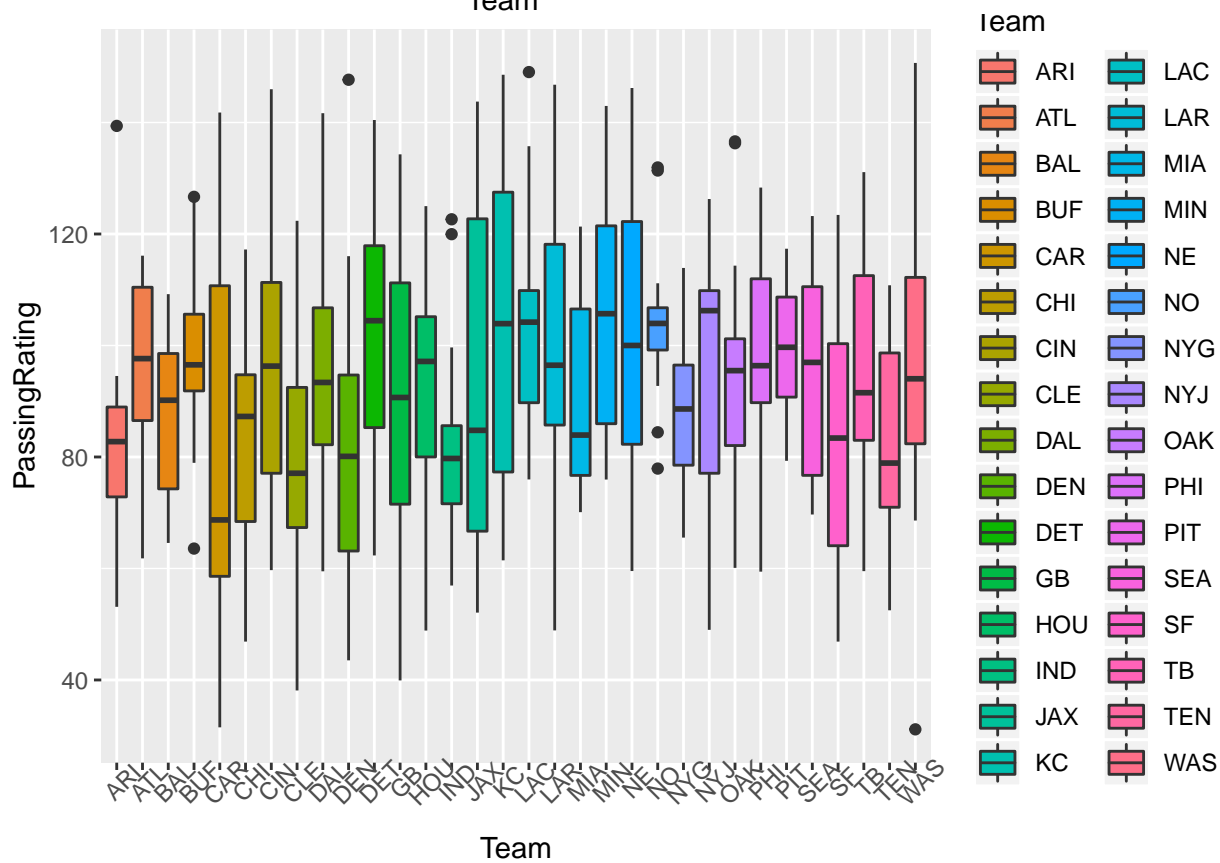
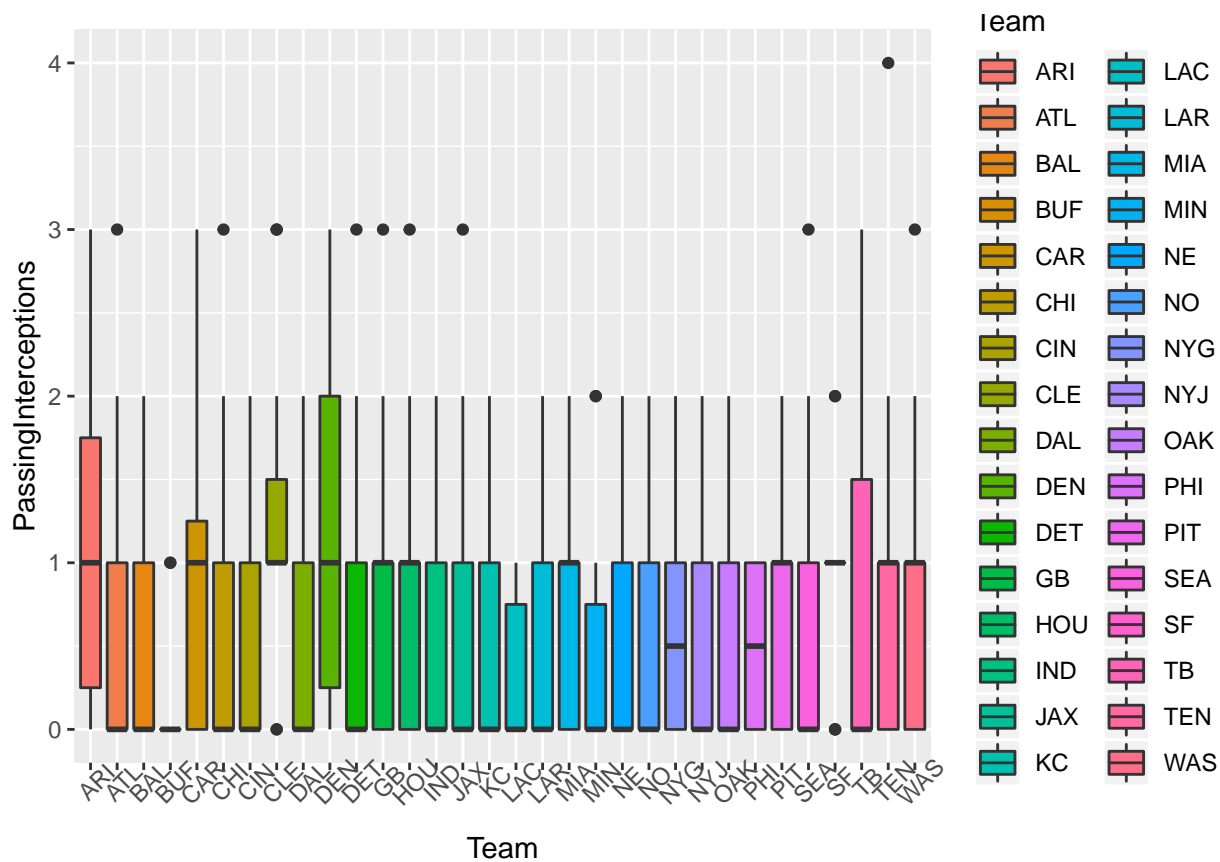
```
for (i in 7:19) {
  ggplotp <- QBCrossSectional %>% ggplot(aes_string(y=names(QBCrossSectional[i]),x="Team",fill="Team"))
    geom_boxplot()+
    xlab("Team")+ylab(names(QBCrossSectional[i]))+
    theme(axis.text.x = element_text(angle=45))

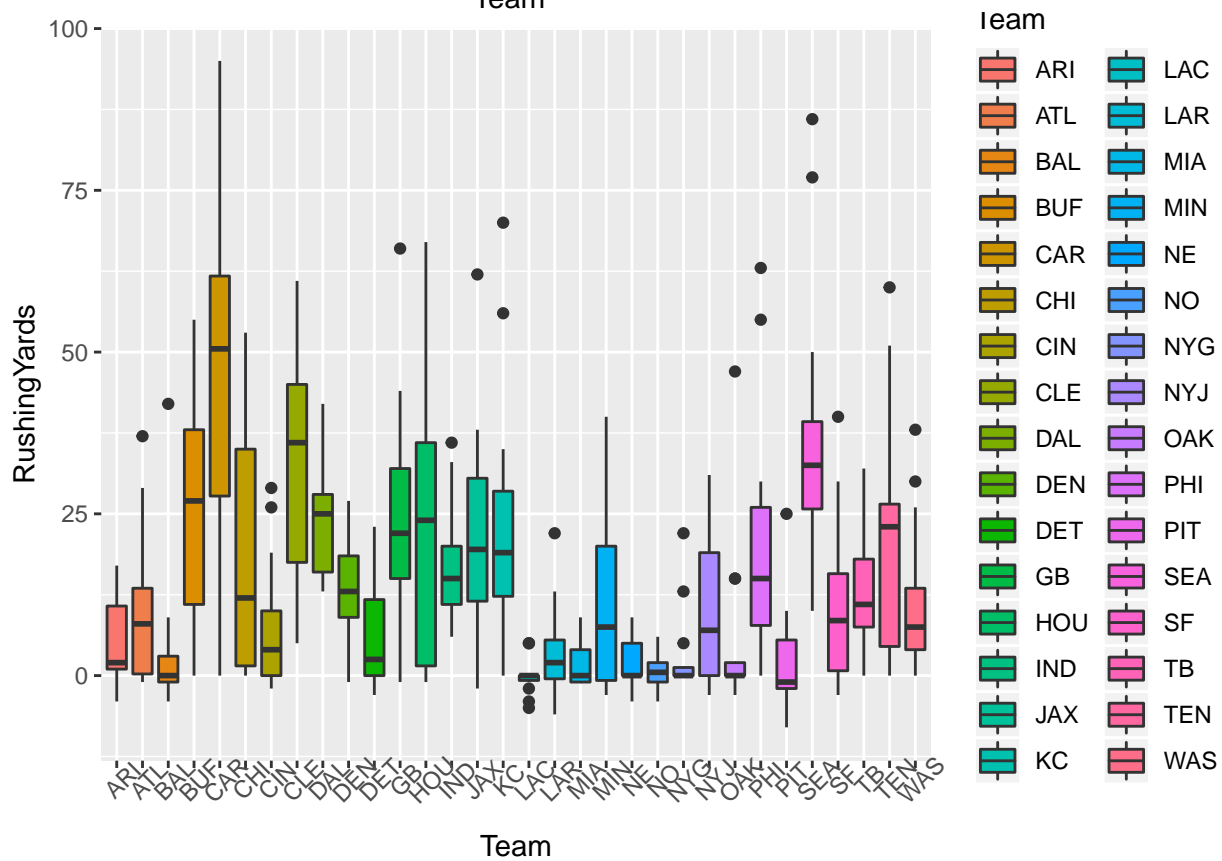
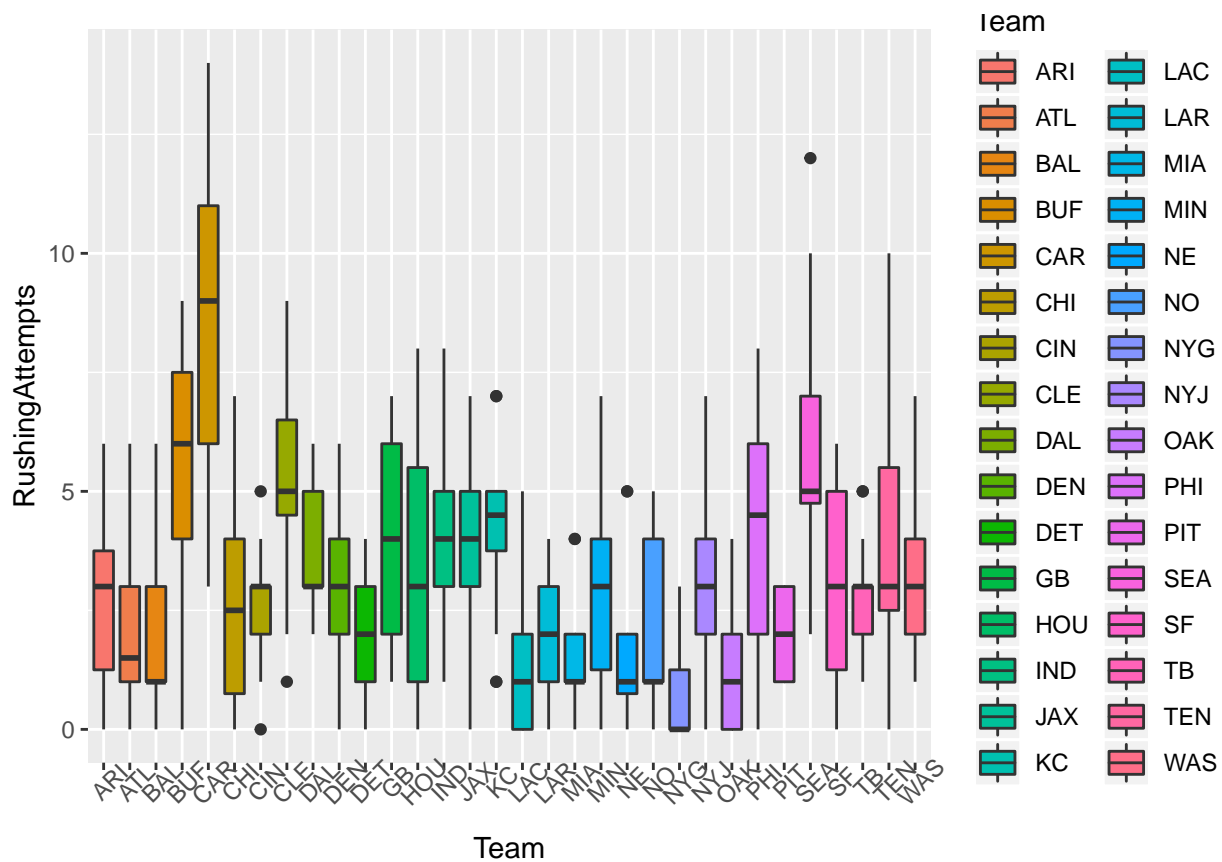
  print(ggplotp)
}
```

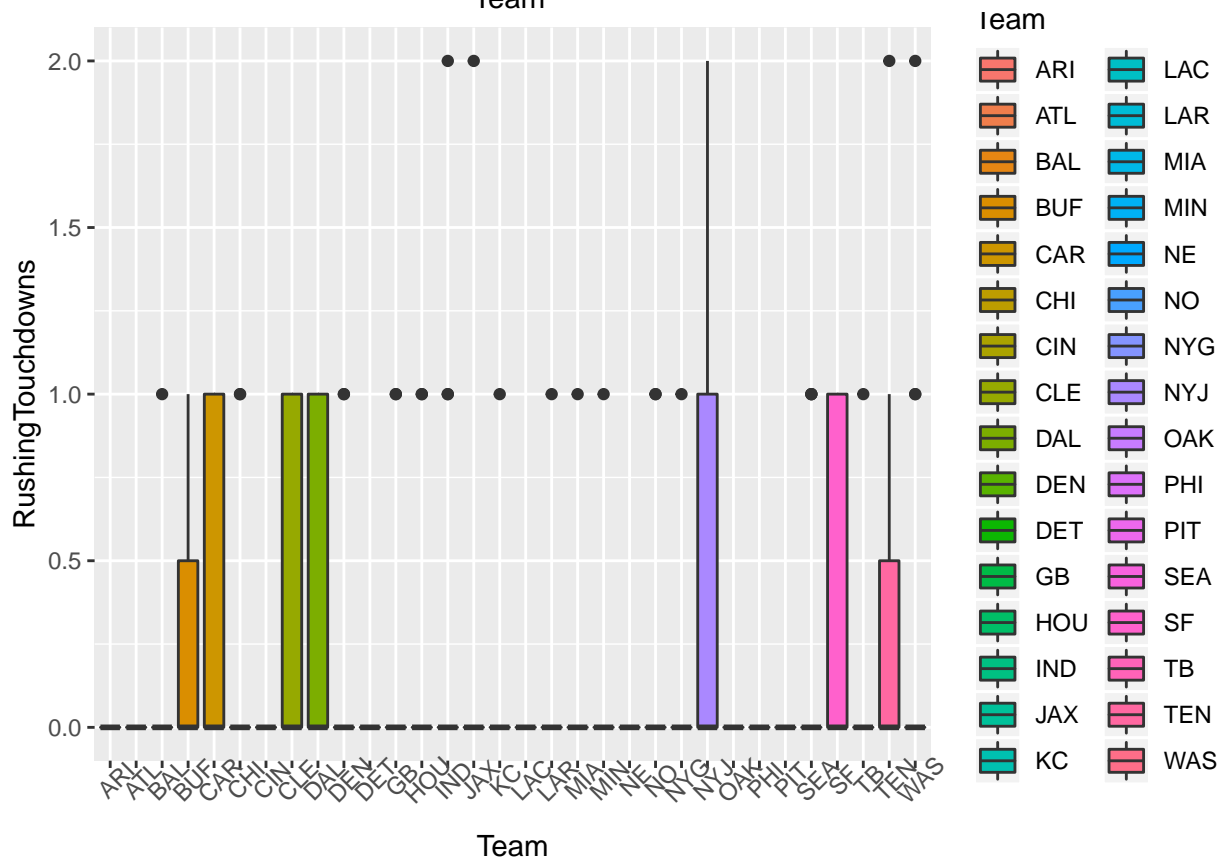
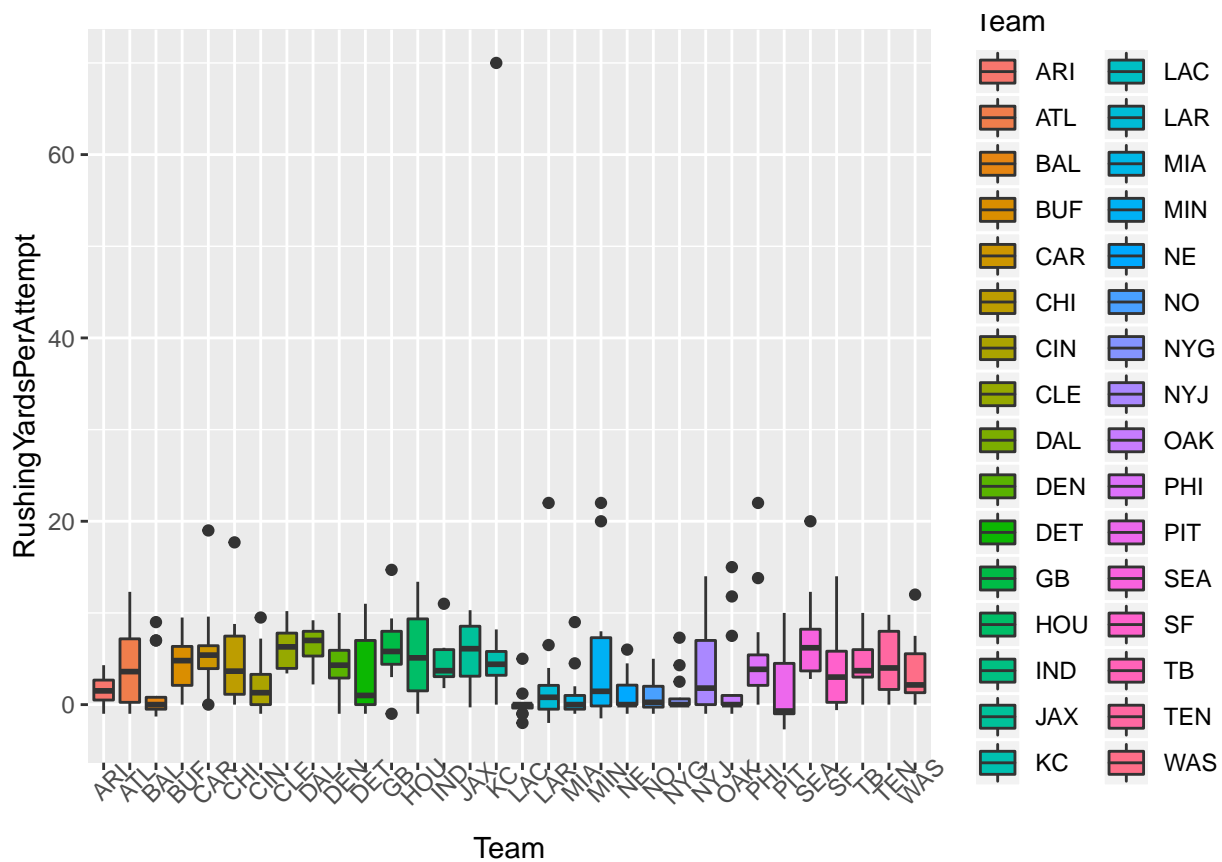


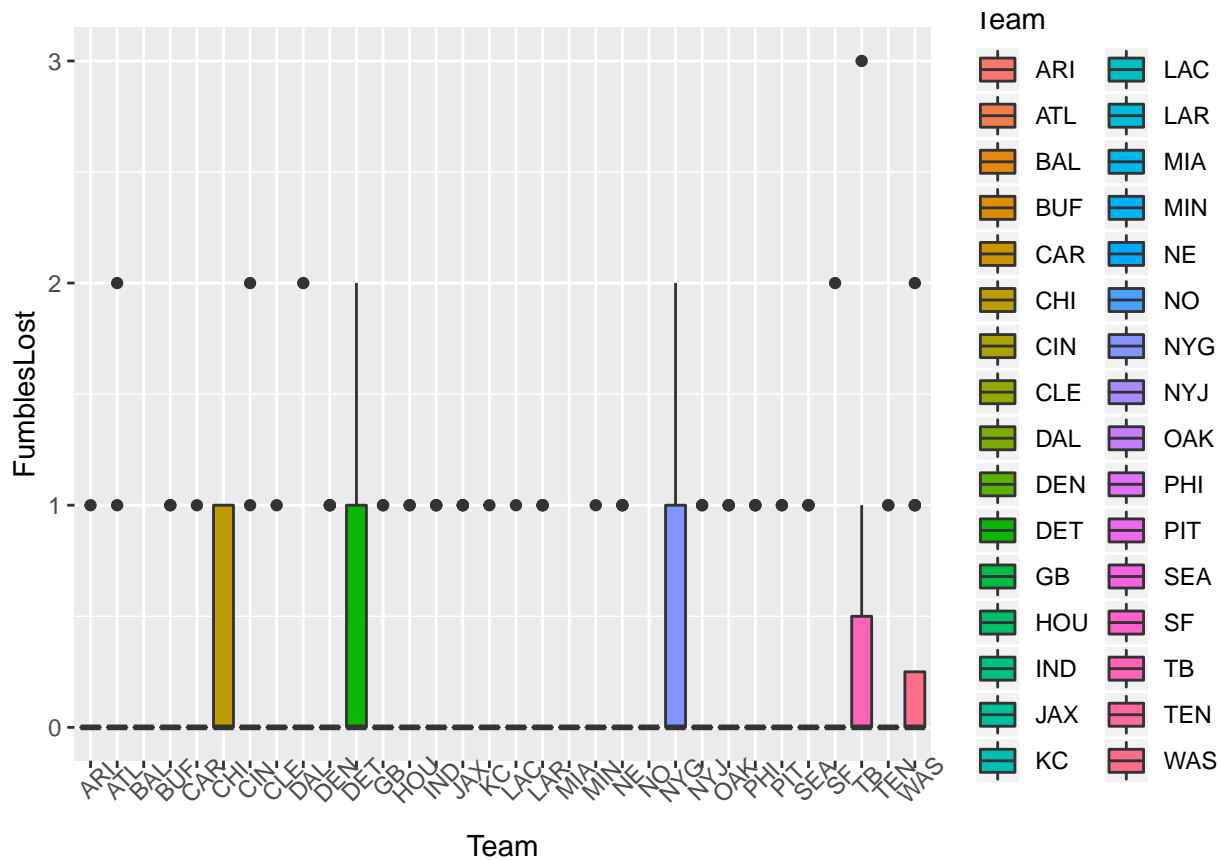












## 1.6: Relationships

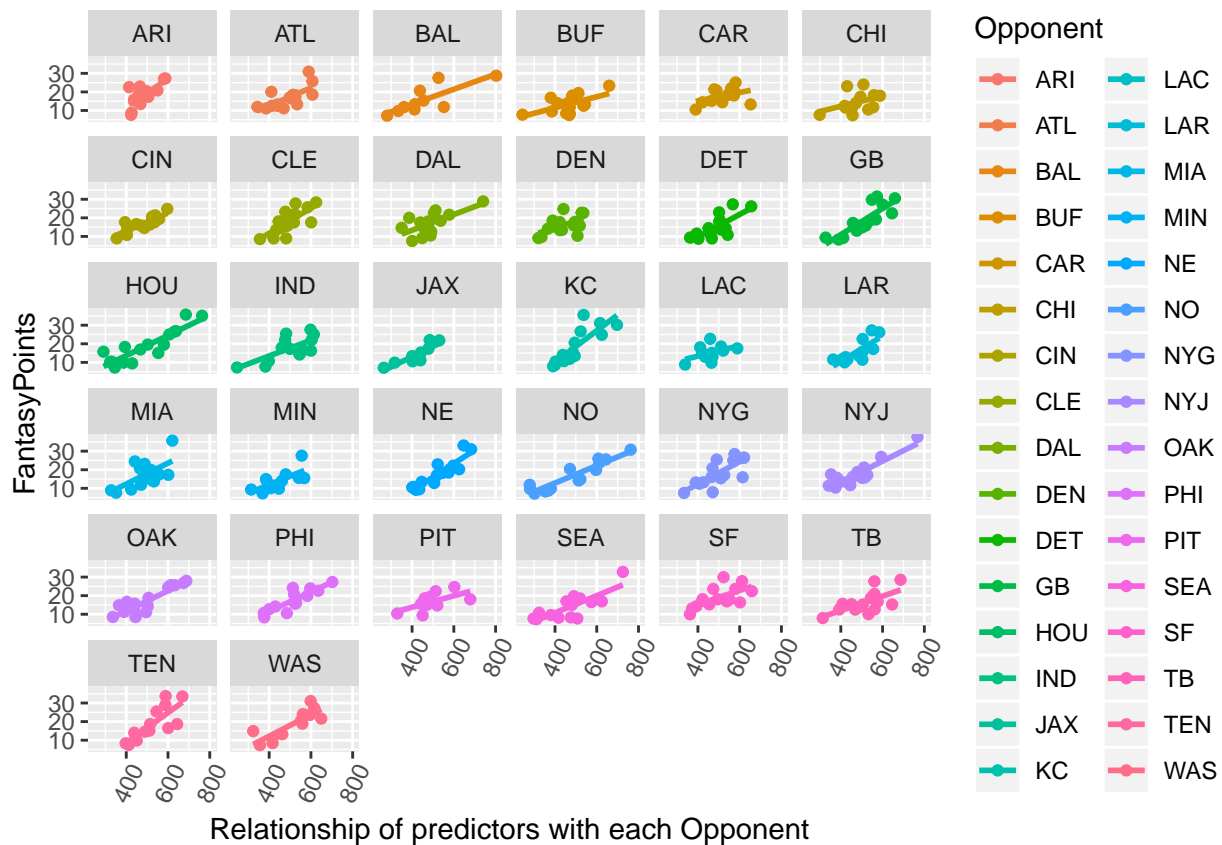
### 1.6.1 Relationship between FantasyPoints with all predictors and taking Opponent into consideration

```
attach(QBCrossSectional)

## The following object is masked from package:ggplot2:
##
## Position

QBCrossSectional %>% ggplot(aes(y=FantasyPoints,x=PassingCompletions+PassingAttempts+
    PassingCompletionPercentage+
    PassingYards+PassingYardsPerAttempt+
    PassingTouchdowns+PassingInterceptions+
    PassingRating+RushingAttempts+RushingYards+
    RushingYardsPerAttempt+RushingTouchdowns+FumblesLost,
    color=Opponent))+
  xlab("Relationship of predictors with each Opponent")+
  geom_point()+
  geom_smooth(method="lm",se=F)+
  theme(axis.text.x = element_text(angle=65, vjust=0.6))+facet_wrap(~Opponent)
```





### 1.6.2 Relationship between FantasyPoints with all predictors and taking Home turf into consideration

```
attach(QBCrossSectional)
```

```
## The following objects are masked from QBCrossSectional (pos = 3):
```

```
##
```

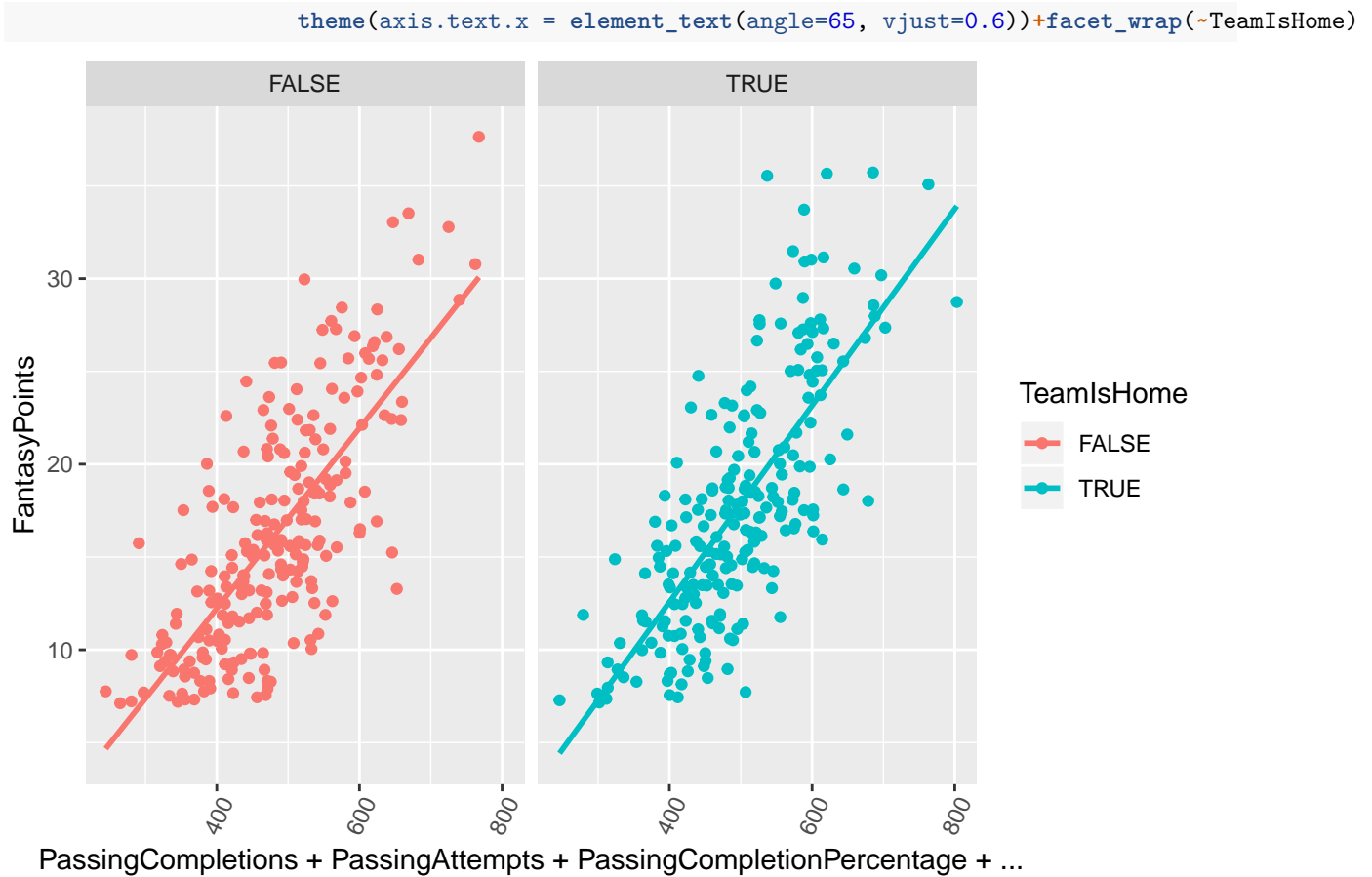
```
## FantasyPoints, FumblesLost, GameDate, Opponent,
## PassingAttempts, PassingCompletionPercentage,
## PassingCompletions, PassingInterceptions, PassingRating,
## PassingTouchdowns, PassingYards, PassingYardsPerAttempt,
## PlayerID, Position, RushingAttempts, RushingTouchdowns,
## RushingYards, RushingYardsPerAttempt, Team, TeamIsHome, Week
```

```
## The following object is masked from package:ggplot2:
```

```
##
```

```
## Position
```

```
QBCrossSectional %>% ggplot(aes(y=FantasyPoints,x=PassingCompletions+PassingAttempts+
    PassingCompletionPercentage+
    PassingYards+PassingYardsPerAttempt+
    PassingTouchdowns+PassingInterceptions+
    PassingRating+RushingAttempts+RushingYards+
    RushingYardsPerAttempt+RushingTouchdowns+FumblesLost,
    color=TeamIsHome))+
    geom_point()+
    geom_smooth(method="lm",se=F)+
```

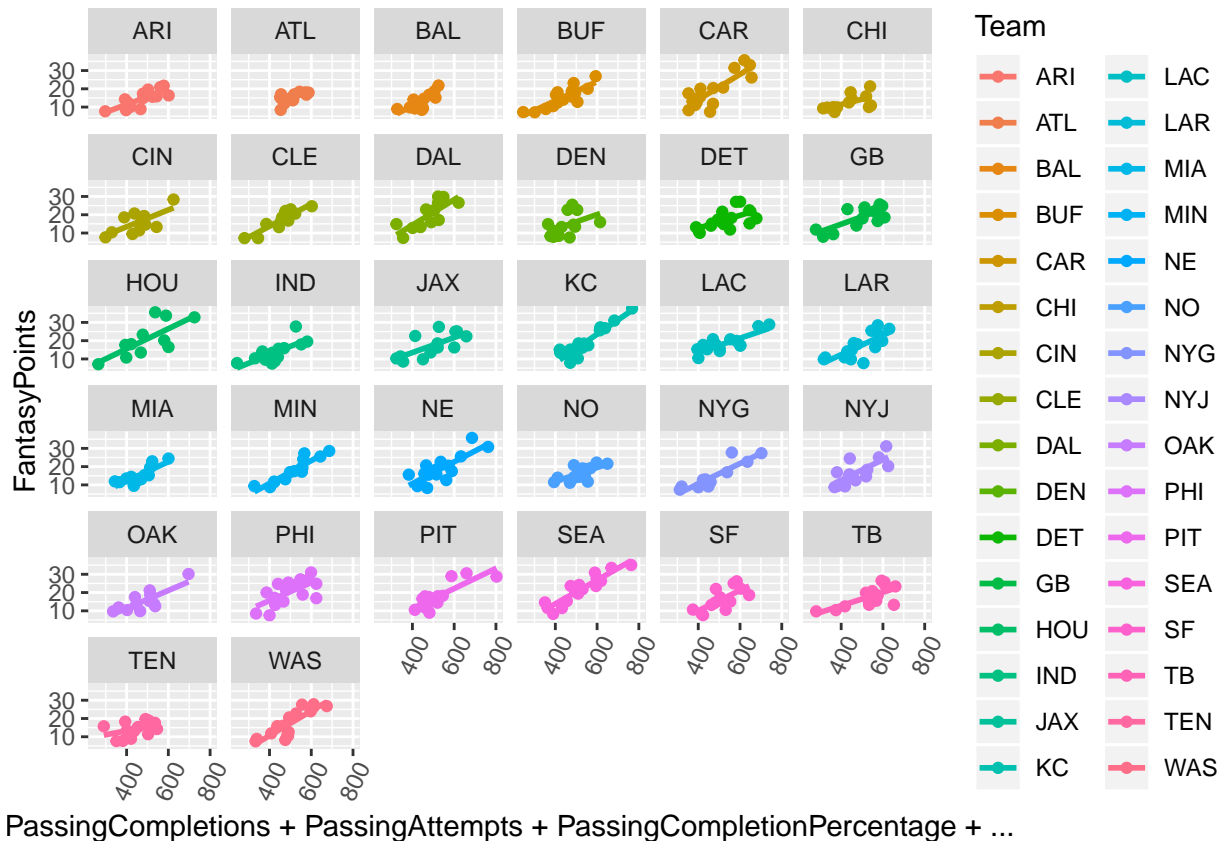


### 1.6.3 Relationship between FanatasyPoints with all predictors per team

```
attach(QBCrossSectional)
```

```
## The following objects are masked from QBCrossSectional (pos = 3):
##
##   FantasyPoints, FumblesLost, GameDate, Opponent,
##   PassingAttempts, PassingCompletionPercentage,
##   PassingCompletions, PassingInterceptions, PassingRating,
##   PassingTouchdowns, PassingYards, PassingYardsPerAttempt,
##   PlayerID, Position, RushingAttempts, RushingTouchdowns,
##   RushingYards, RushingYardsPerAttempt, Team, TeamIsHome, Week
## The following objects are masked from QBCrossSectional (pos = 4):
##
##   FantasyPoints, FumblesLost, GameDate, Opponent,
##   PassingAttempts, PassingCompletionPercentage,
##   PassingCompletions, PassingInterceptions, PassingRating,
##   PassingTouchdowns, PassingYards, PassingYardsPerAttempt,
##   PlayerID, Position, RushingAttempts, RushingTouchdowns,
##   RushingYards, RushingYardsPerAttempt, Team, TeamIsHome, Week
## The following object is masked from package:ggplot2:
##
##   Position
```

```
QBCrossSectional %>% ggplot(aes(y=FantasyPoints,x=PassingCompletions+PassingAttempts+
    PassingCompletionPercentage+
    PassingYards+PassingYardsPerAttempt+
    PassingTouchdowns+PassingInterceptions+
    PassingRating+RushingAttempts+
    RushingYards+RushingYardsPerAttempt+
    RushingTouchdowns+FumblesLost,color=Team)) +
    geom_point()+
    geom_smooth(method="lm",se=F)+
    theme(axis.text.x = element_text(angle=65, vjust=0.6))+facet_wrap(~Team)
```



#### 1.6.4 Relationship between FantasyPoints and Individual predictors

```
for (i in 7:19) {
  linear_plot <- QBCrossSectional %>%
    ggplot(aes_string(y="FantasyPoints",x=names(QBCrossSectional[i]))) +
    geom_point()+geom_smooth(method="lm",se=F)+
    theme(axis.text.x = element_text(angle=65, vjust=0.6))+
    theme_economist()
  print(linear_plot)
}
```

