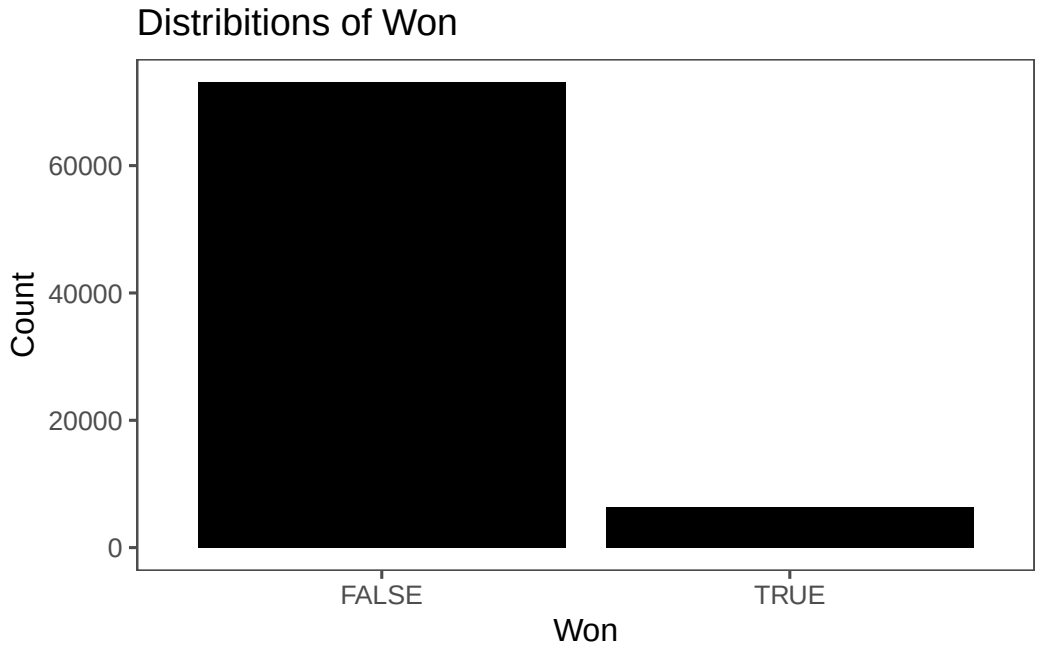
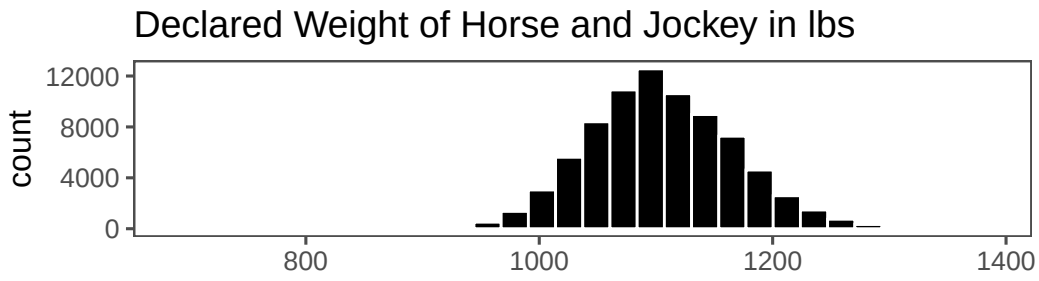
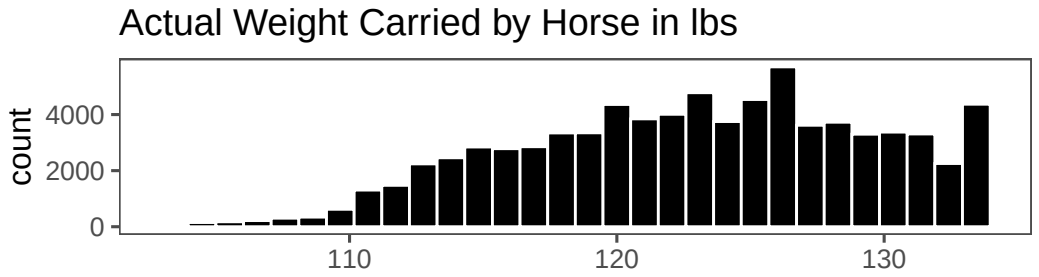
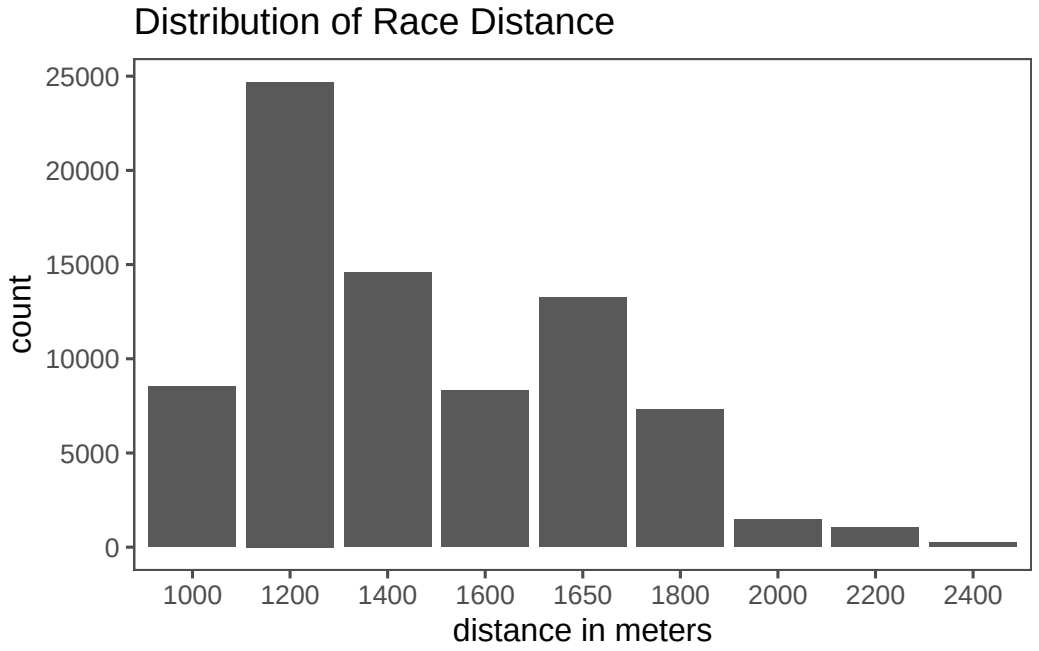
EDA



The average number of horses per race is 12.65. On average the chance odds against winning are $\frac{1-p}{p} = \frac{11.65}{1}$, so $11.65 : 1$. Even though there is ≈ 0.079 chance of winning, with 806 wins in our 10,000 training data we satisfy the one in ten rule.

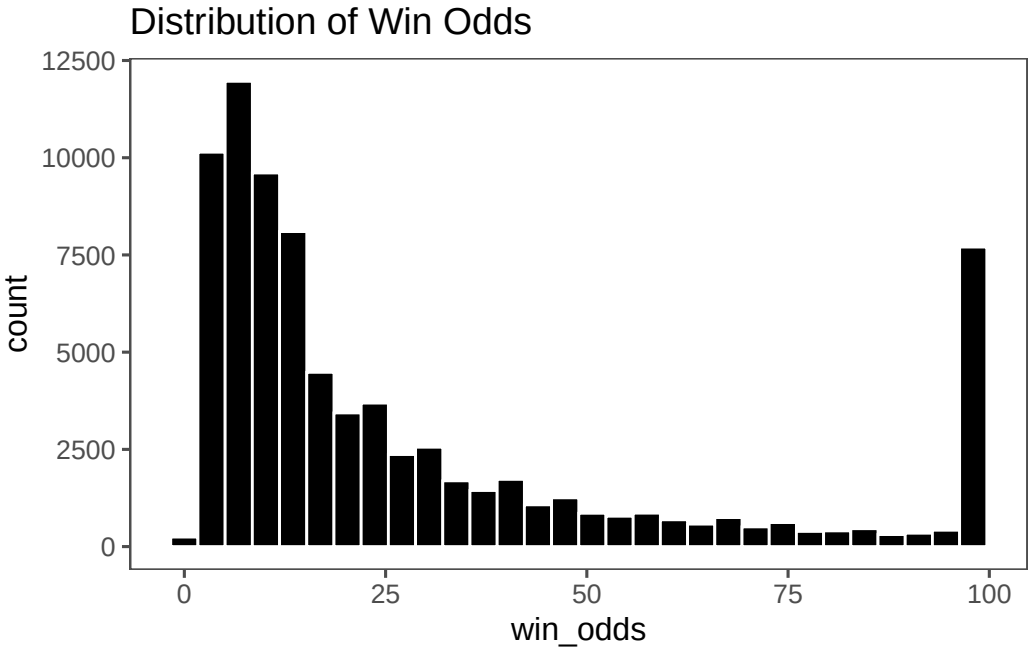


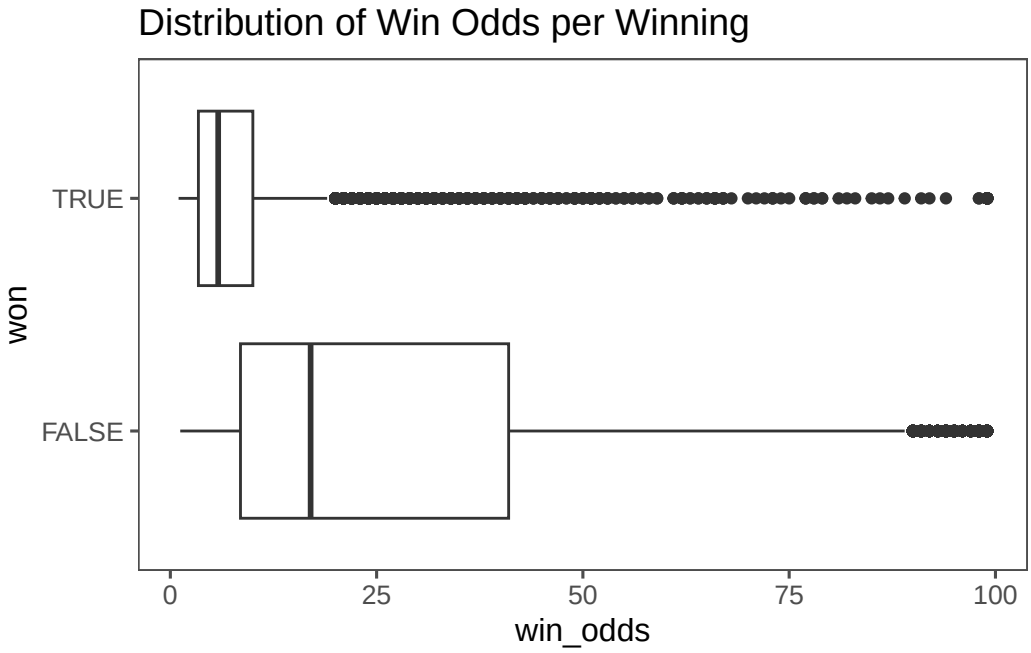


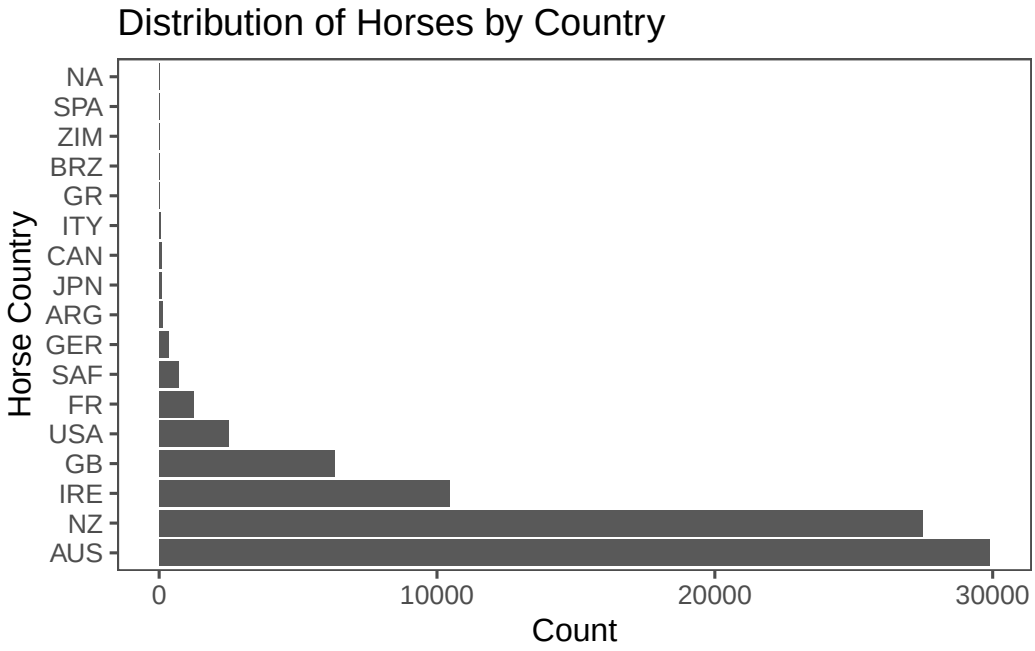


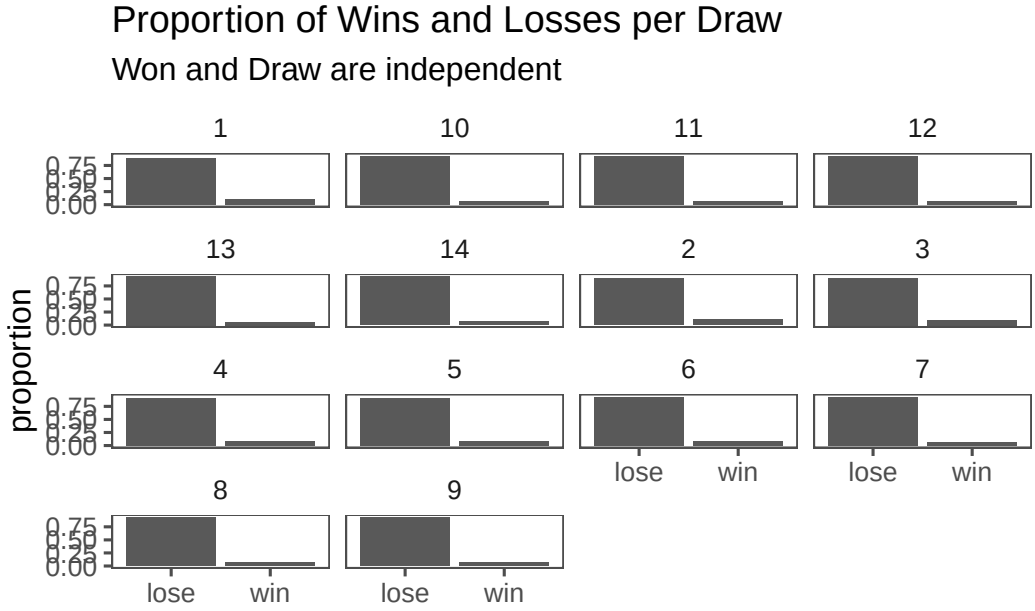

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  10.00   60.00   60.00   61.03   60.00  138.00 

10% 20% 80% 90% 
```

```
52  60  60  72
```


Distribution of Win Odds
12500
10000
7500
5000
2500
0
count
0
25
50
75
100
win_odds


Distribution of Win Odds per Winning
TRUE
FALSE
won
0
25
50
75
100
win_odds


Distribution of Horses by Country
Horse Country
NA
SPA
ZIM
BRZ
GR
ITY
CAN
JPN
ARG
GER
SAF
FR
USA
GB
IRE
NZ
AUS
0
10000
20000
30000
Count


Proportion of Wins and Losses per Draw
Won and Draw are independent
proportion
1
10
11
12
13
14
2
3
4
5
6
7
8
9
0.75
0.50
0.25
0.00
lose
win

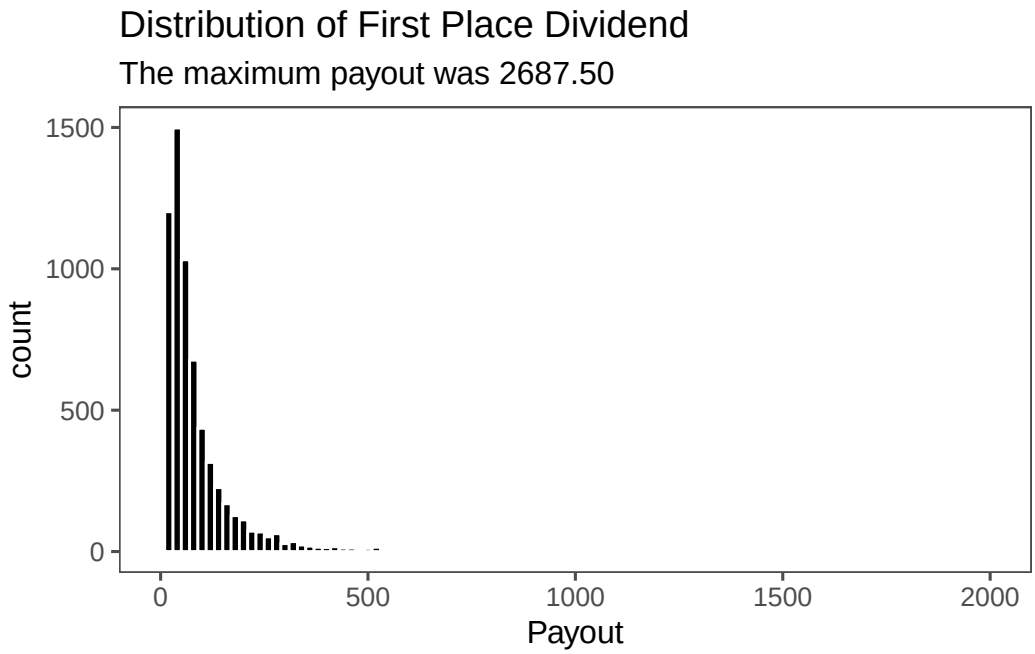

Distribution of First Place Dividend
The maximum payout was 2687.50
count
1500
1000
500
0
0
500
1000
1500
2000
Payout

Table 1: Number of Levels per Character Variable

Variable	Number of Levels
race_id	6348
horse_id	4405
horse_gear	822
jockey_id	186
trainer_id	176
horse_ratings	31
horse_country	17
draw	15
position_sec4	15
position_sec5	15
position_sec6	15
place_combination3	15
horse_no	14
result	14
position_sec1	14
position_sec2	14
position_sec3	14
place_combination1	14
place_combination2	14
race_no	11
place_combination4	11
horse_type	10
going	10
race_class	10
distance	9
config	6
venue	2
surface	2