BIG DATA AND DATA MINING PROJECT:
2019 UK ROAD ACCIDENT REPORT

MSc in Artificial Intelligence and Data Science

OSAGIE ELLIOT AIBANGBEE

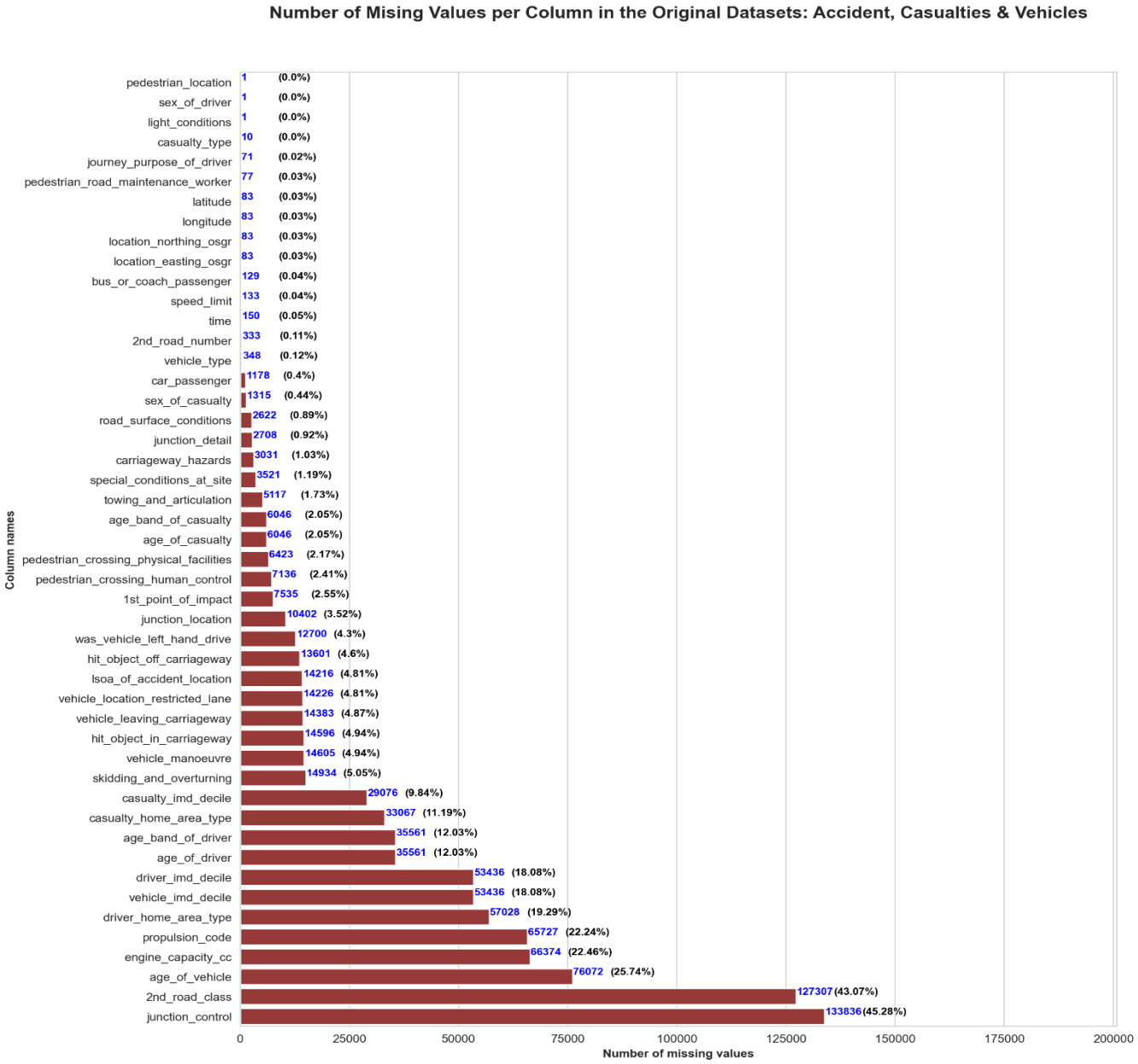
STUDENT NUMBER: 20211557

APRIL 2022

INTRODUCTION

The 3 datasets were merged on their common identifier, 'accident_index', and the result was a table having 2 95,579 rows and 69 columns. My goal is to extract sufficient knowledge from the data through analyses and u se this knowledge to build a predictive model that can assist government officials in making precautionary po licies to prevent enabling conditions for road accidents.

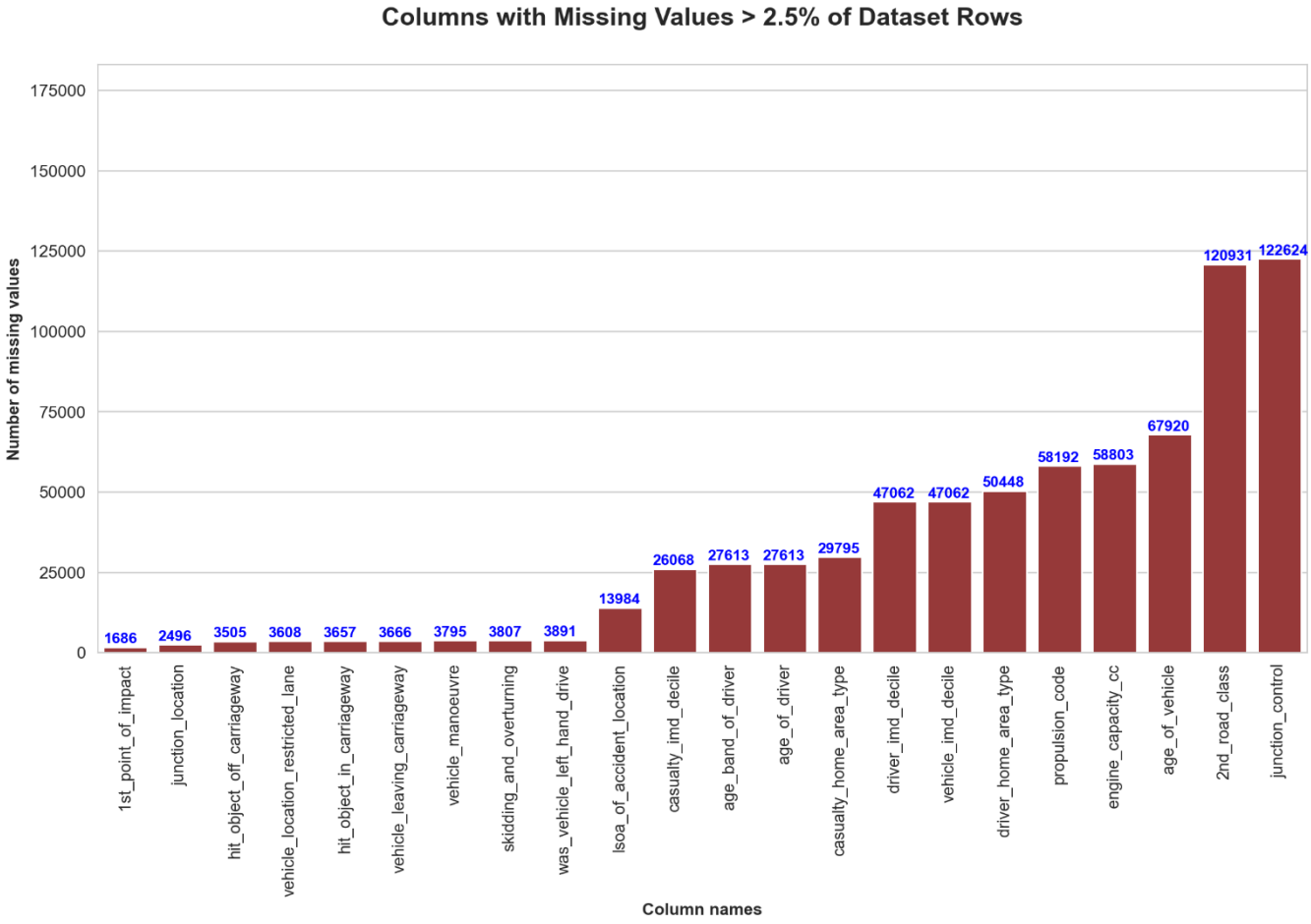
DATA CLEANING



i. HANDLING MISSING VALUES

1. DROPPING SOME MISSING ENTRIES

All missing values in columns with less than or equal to 2.5% of total entries were dropped.



Most nulls were replaced with modal values for each unique category of some correlated variables.

a) *Using the Top 2 or 3 Most Correlated Features as Pivot for Modal/Median Values*: Most of the above columns were imputed following this pattern.
b) In the 2nd_road_class column, nulls were replaced with 6, for 'unclassified' and for propulsion_code, 0 replaced 'undefined' thereby retaining its numeric datatype.

Retained 93% of the data (after cleaning) for further analysis and prediction.

ii. FEATURE ENGINEERING

Engineered features include:

1. *part_of_day*:
   1 → morning (5am – 11.59am)
   2 → afternoon (12pm – 16.59pm)
   3 → evening (17 – 20.59pm)
   4 → night (21pm – 4.59am)
2. *is_weekend:* Weekend is Saturday and Sunday.
3. *quarter:* yearly quarters
4. *season*: (autumn, spring, summer, and winter).
5. *is_dst* (2019 Daylight Saving Time): 1am March 31 to 2am October 27.
6. *is_offseason (Premier League offseason indicator): 13 May – 8 August 2019*
7. *is_sunrise and is_sunset: sunrise/sunset time boundaries*

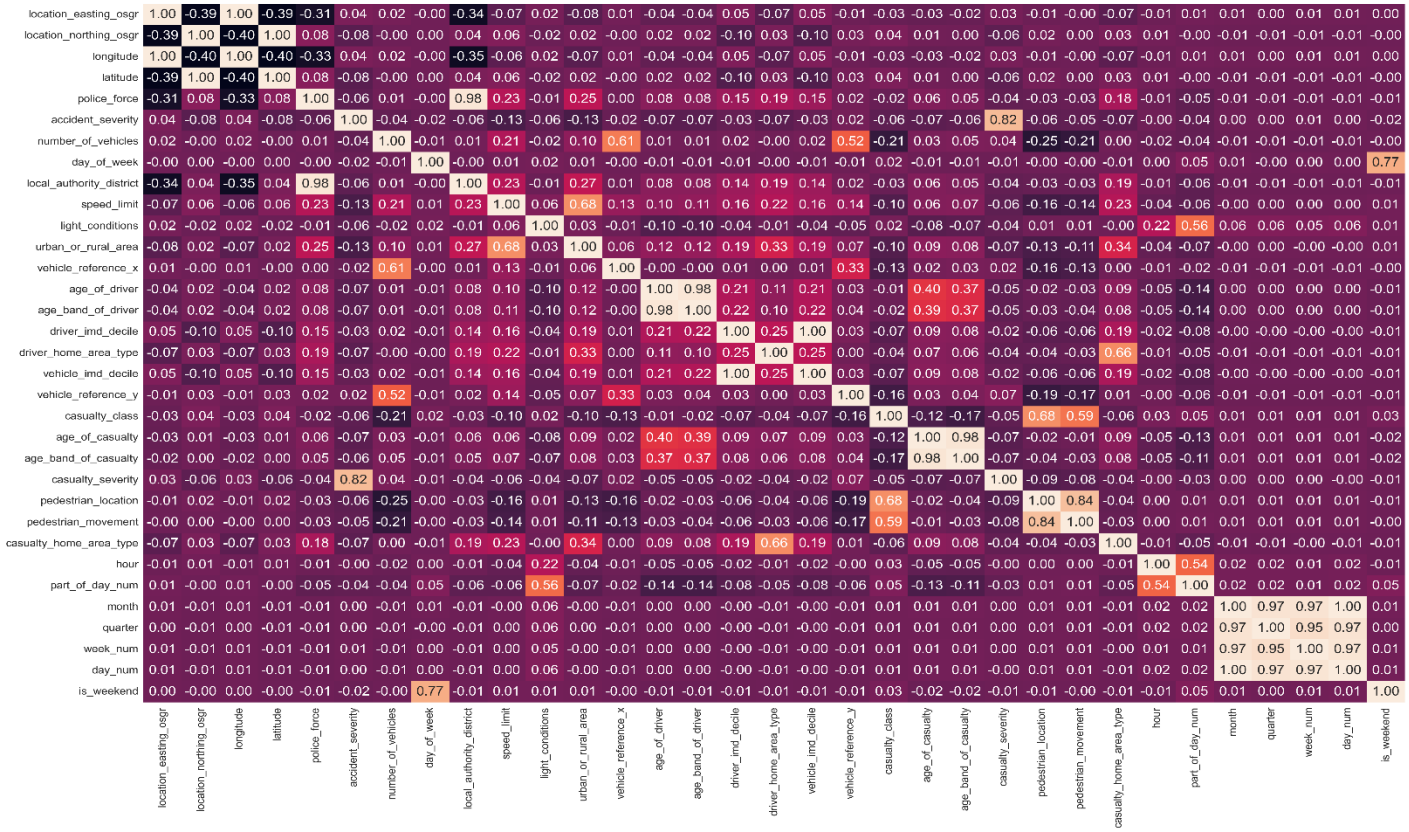
The *time* and *date* columns were dropped to minimize information duplicity.

```
 67  hour                                275520 non-null  int32
 68  minute                              275520 non-null  int32
 69  part_of_day_num                     275520 non-null  float64
 70  month                               275520 non-null  int32
 71  day                                 275520 non-null  int32
 72  quarter                             275520 non-null  int64
 73  month_name                          275520 non-null  object
 74  week_num                            275520 non-null  UInt32
 75  day_num                             275520 non-null  int64
 76  day_name                            275520 non-null  object
 77  is_weekend                          275520 non-null  int64
 78  season                              275520 non-null  object
 79  is_dst                              275520 non-null  int32
 80  is_offseason                        275520 non-null  int32
 81  is_sunrise                          275520 non-null  int64
 82  is_sunset                           275520 non-null  int64
dtypes: UInt32(1), float64(64), int32(6), int64(6), object(6)
memory usage: 167.4+ MB
```

ANALYSIS

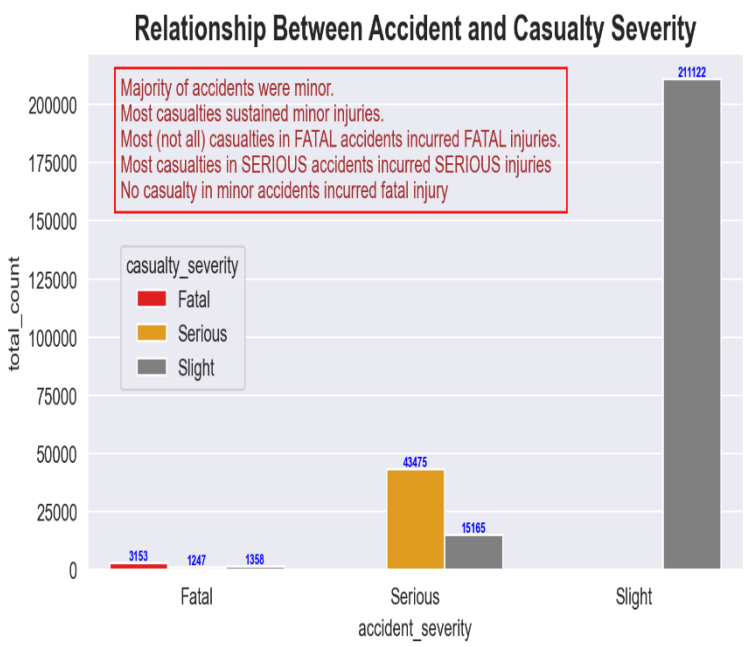
a) CORRELATION (PEARSON'S) ANALYSIS

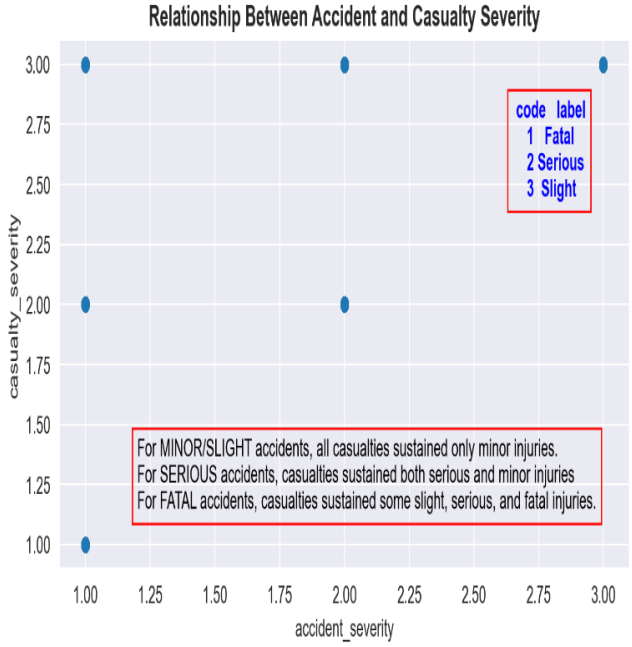
Correlated variables are those with absolute correlation values >= 0.5 (50%)

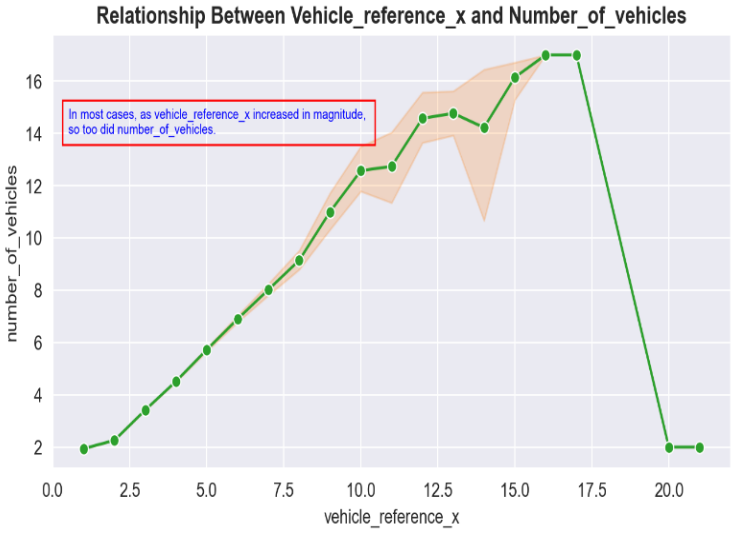


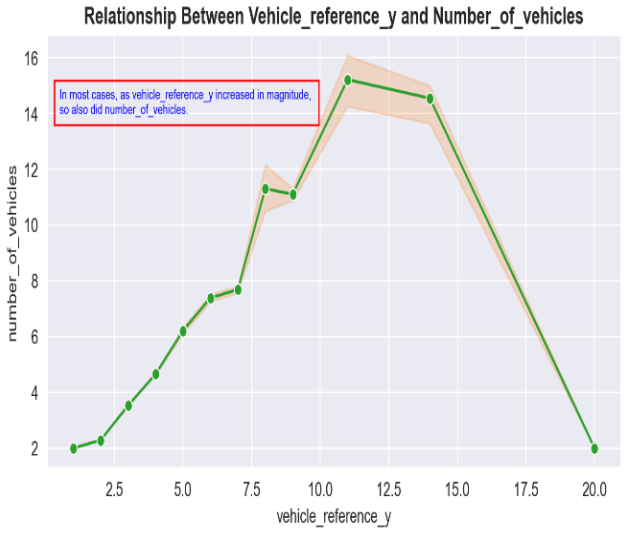
b) AUTO-CORRELATION ANALYSIS

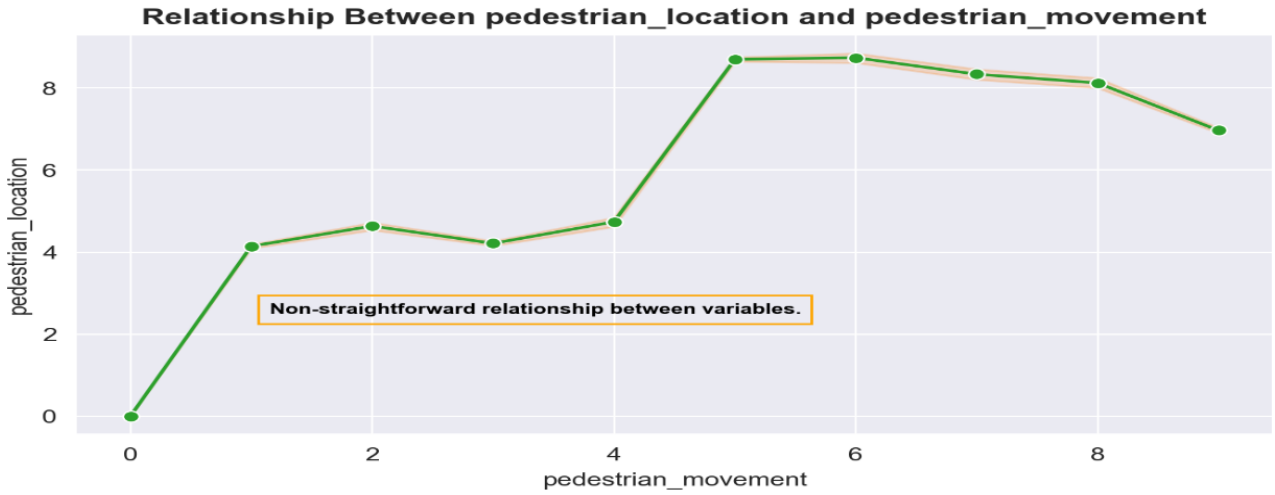
Below variables convey the same information.


Relationship Between Accident and Casualty Severity
Majority of accidents were minor.
Most casualties sustained minor injuries.
Most (not all) casualties in FATAL accidents incurred FATAL injuries.
Most casualties in SERIOUS accidents incurred SERIOUS injuries
No casualty in minor accidents incurred fatal injury
casualty_severity
Fatal
Serious
Slight
total_count
3153
1247
1358
43475
15165
211122
Fatal
Serious
Slight
accident_severity

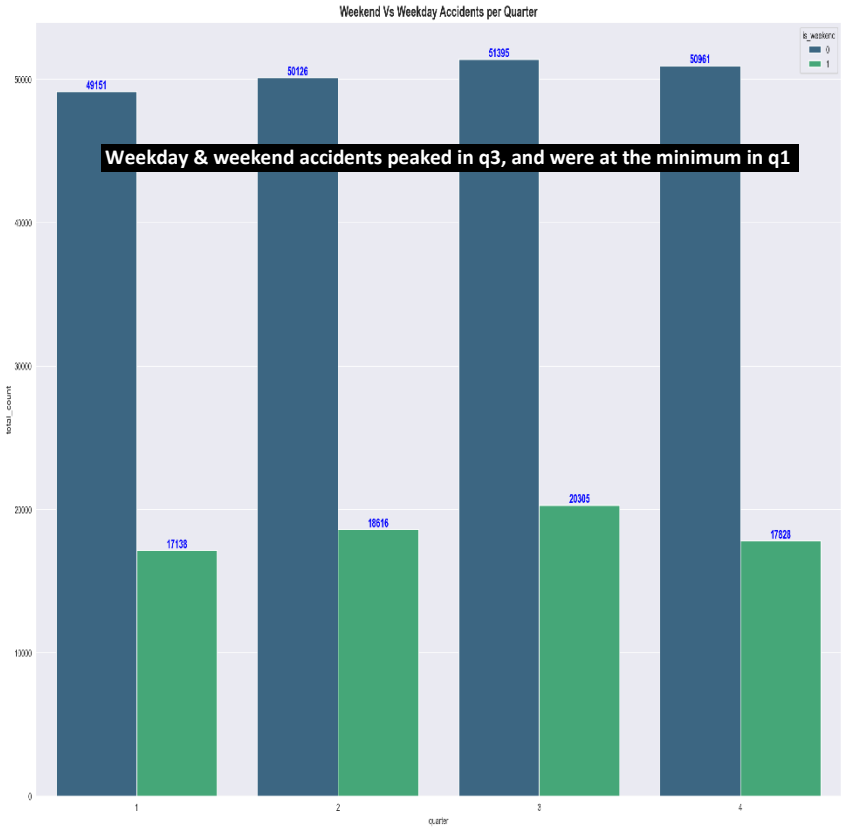

Relationship Between Accident and Casualty Severity
code label
1 Fatal
2 Serious
3 Slight
For MINOR/SLIGHT accidents, all casualties sustained only minor injuries.
For SERIOUS accidents, casualties sustained both serious and minor injuries
For FATAL accidents, casualties sustained some slight, serious, and fatal injuries.
casualty_severity
accident_severity

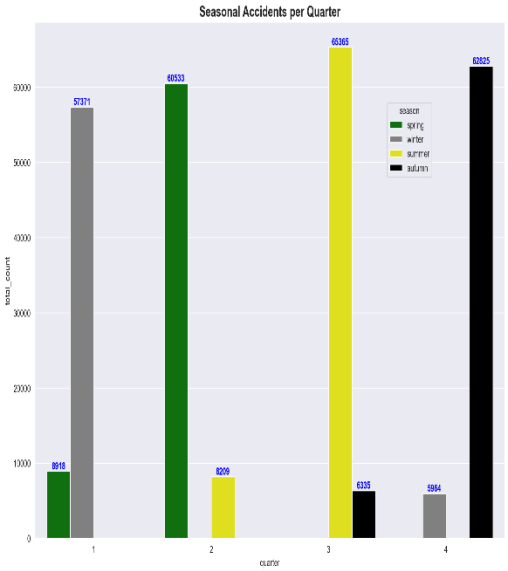

Relationship Between Vehicle_reference_x and Number_of_vehicles
In most cases, as vehicle_reference_x increased in magnitude, so too did number_of_vehicles.
number_of_vehicles
vehicle_reference_x

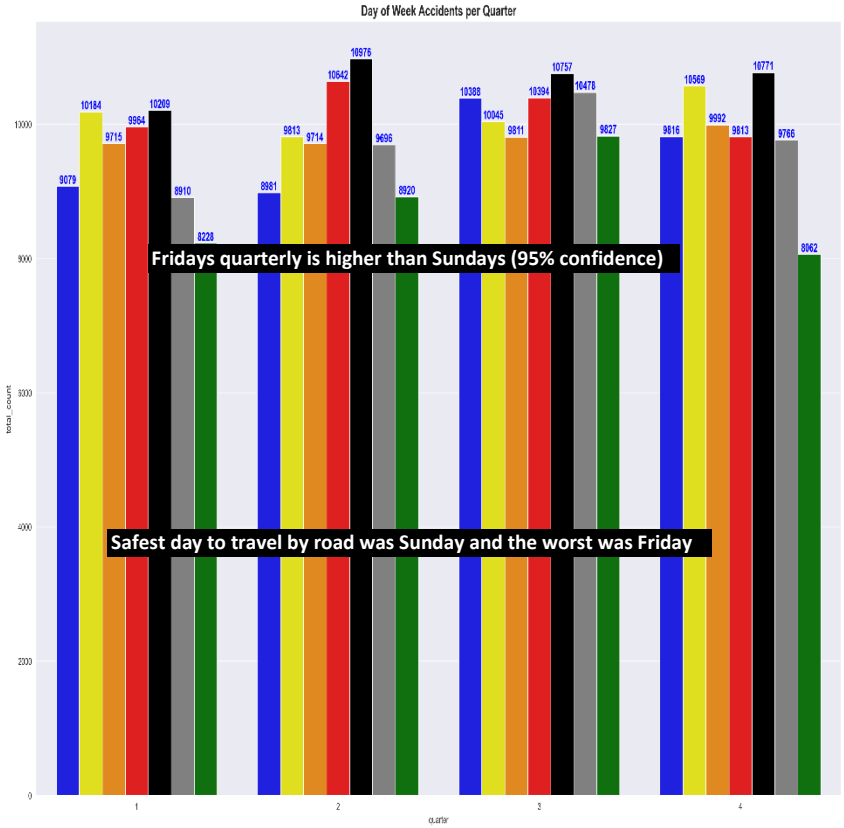

Relationship Between Vehicle_reference_y and Number_of_vehicles
In most cases, as vehicle_reference_y increased in magnitude, so also did number_of_vehicles.
number_of_vehicles
vehicle_reference_y

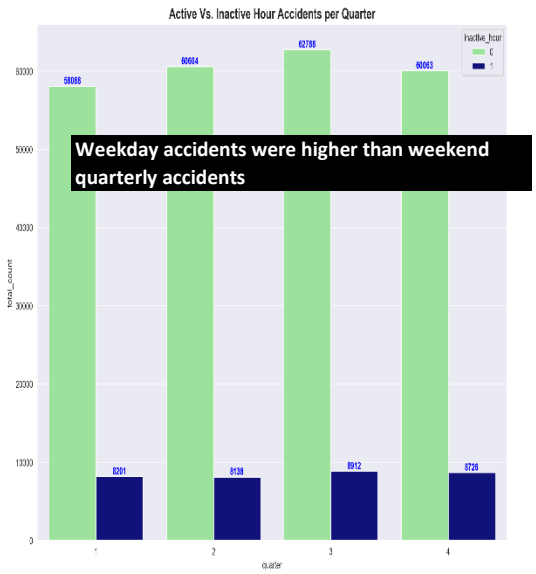

Relationship Between pedestrian_location and pedestrian_movement
Non-straightforward relationship between variables.
pedestrian_location
pedestrian_movement

QUARTERLY OBSERVATIONS

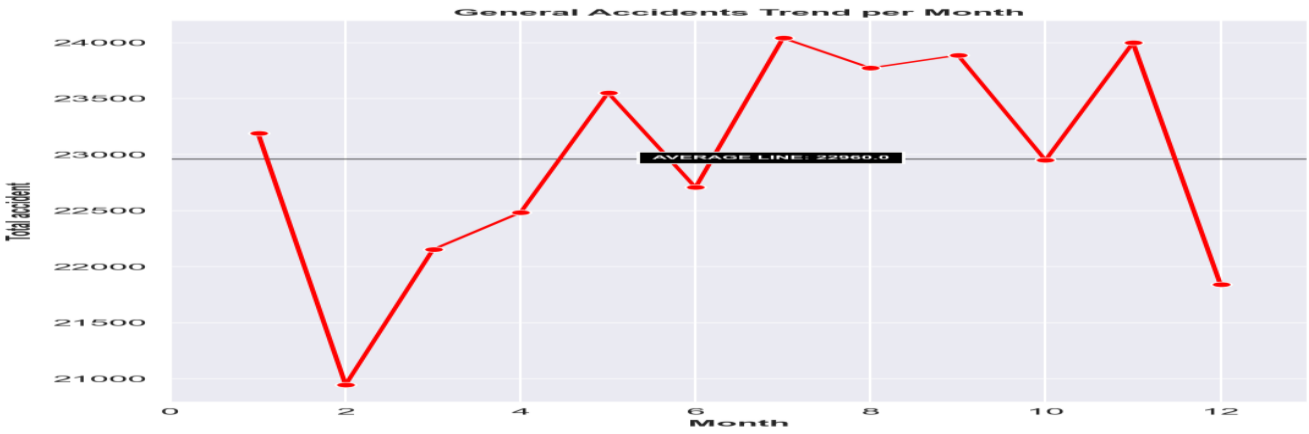


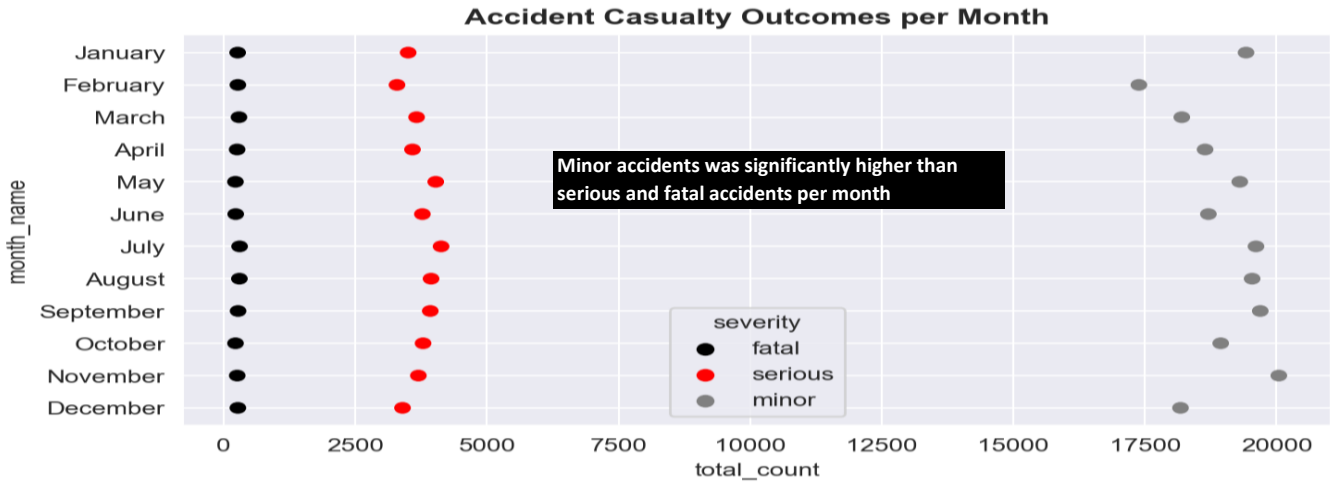


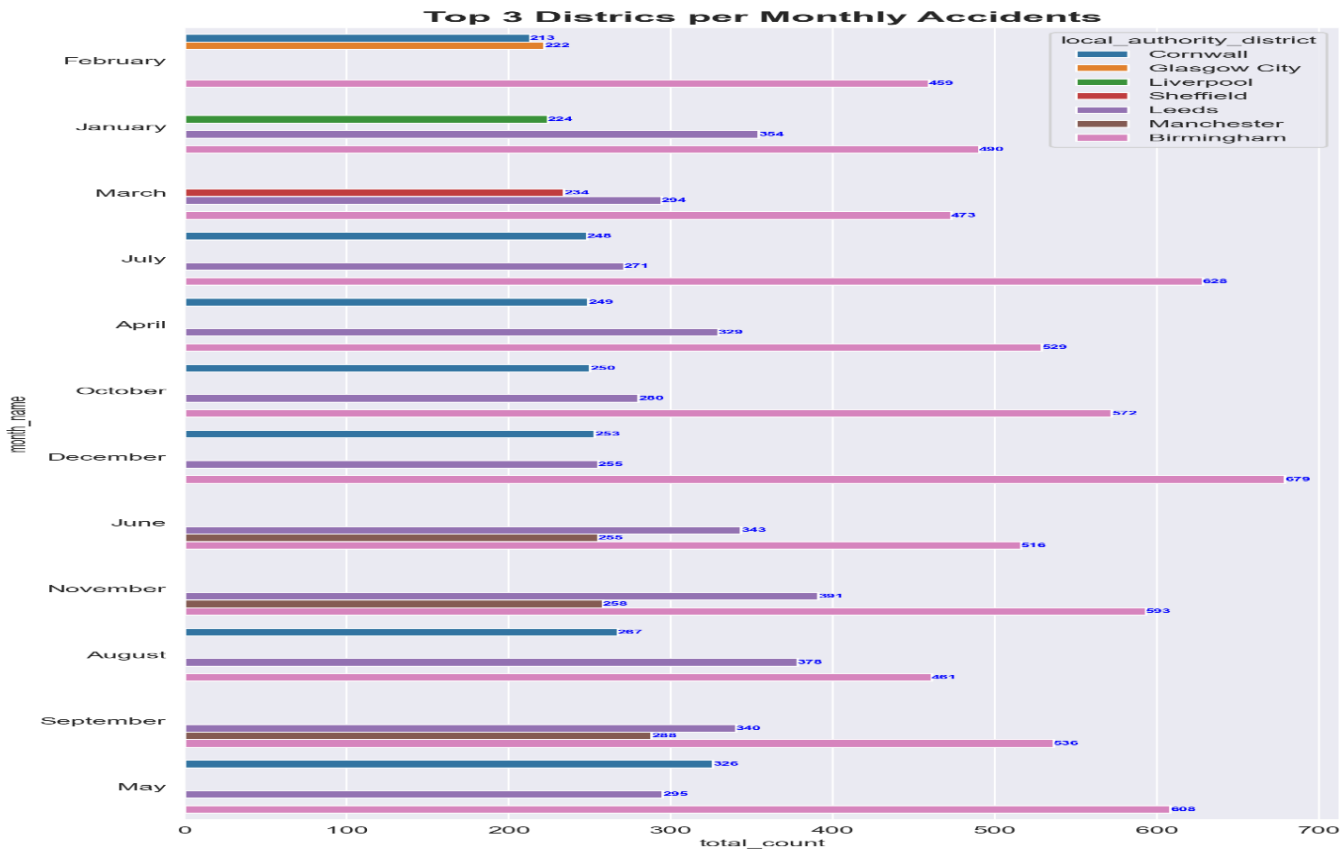


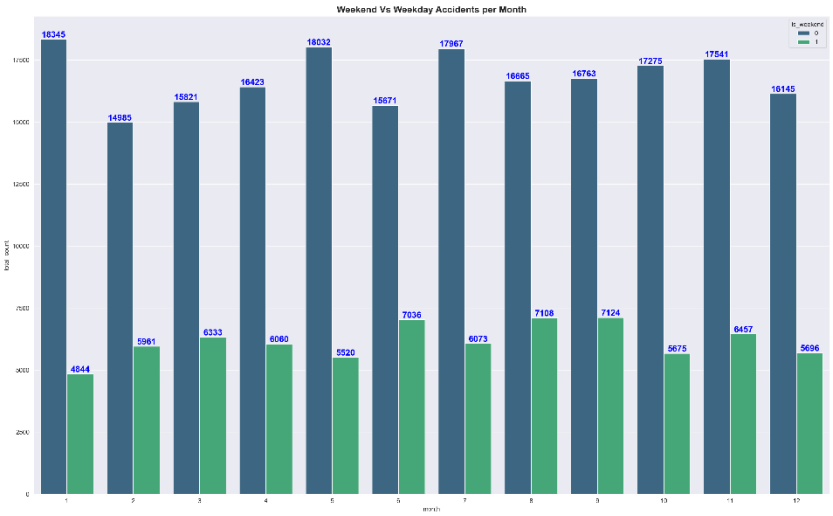


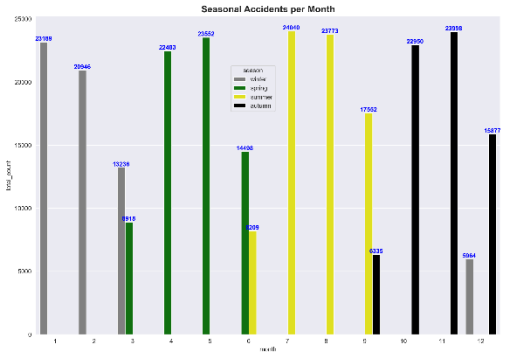
MONTHLY OBSERVATIONS

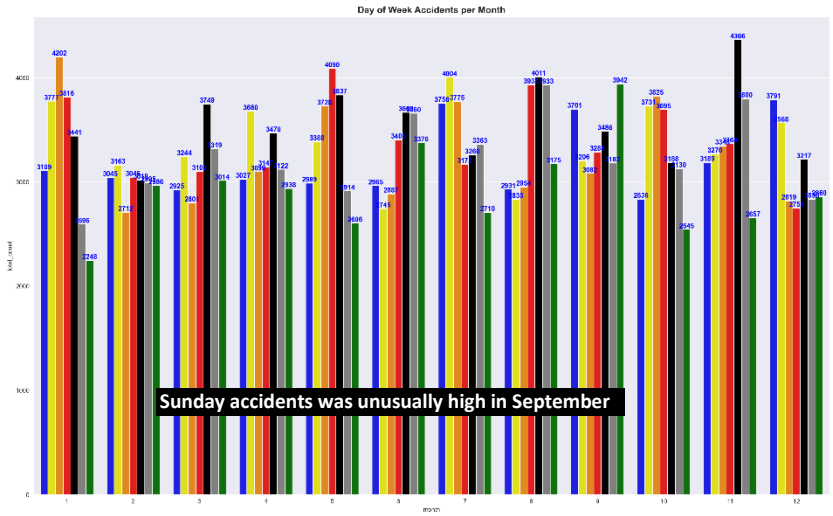


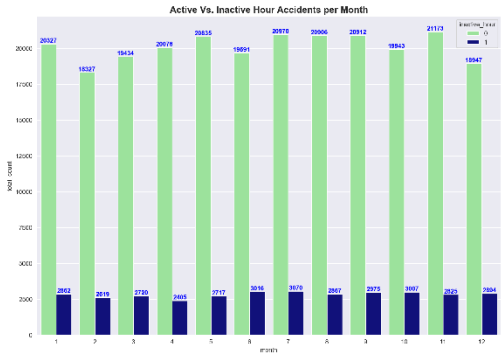




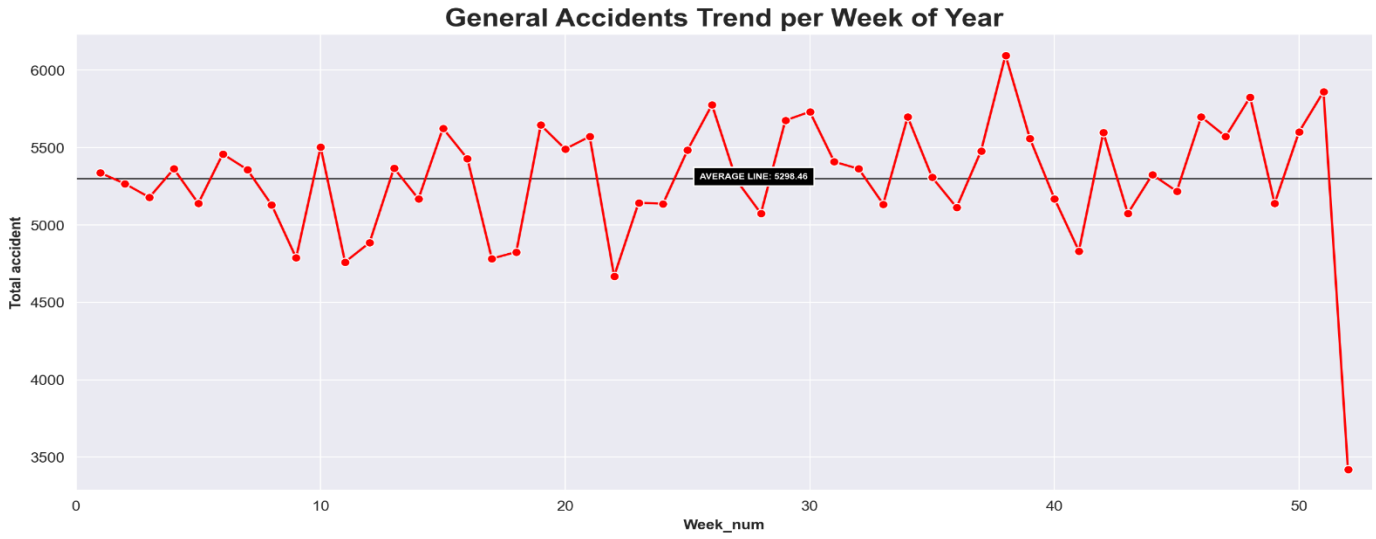


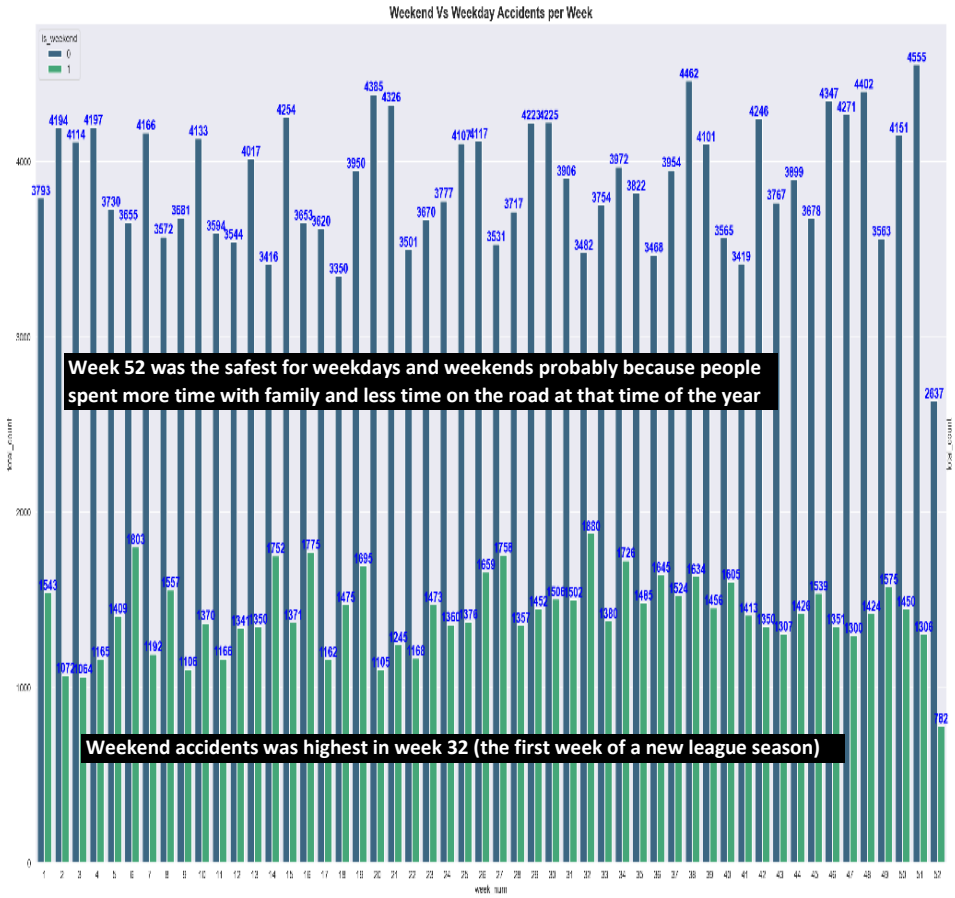


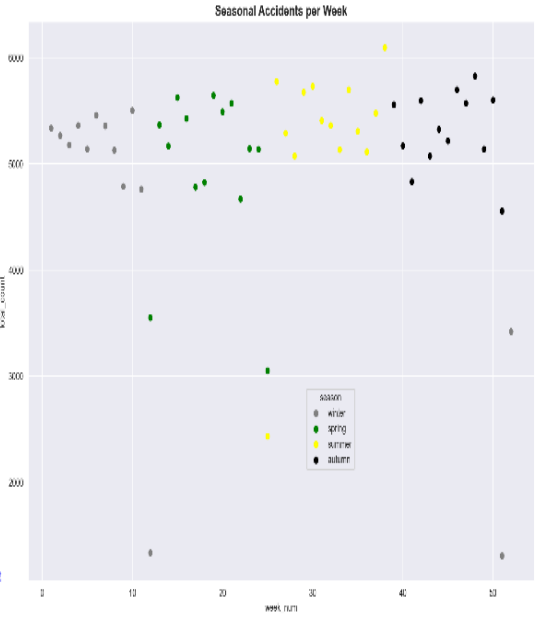


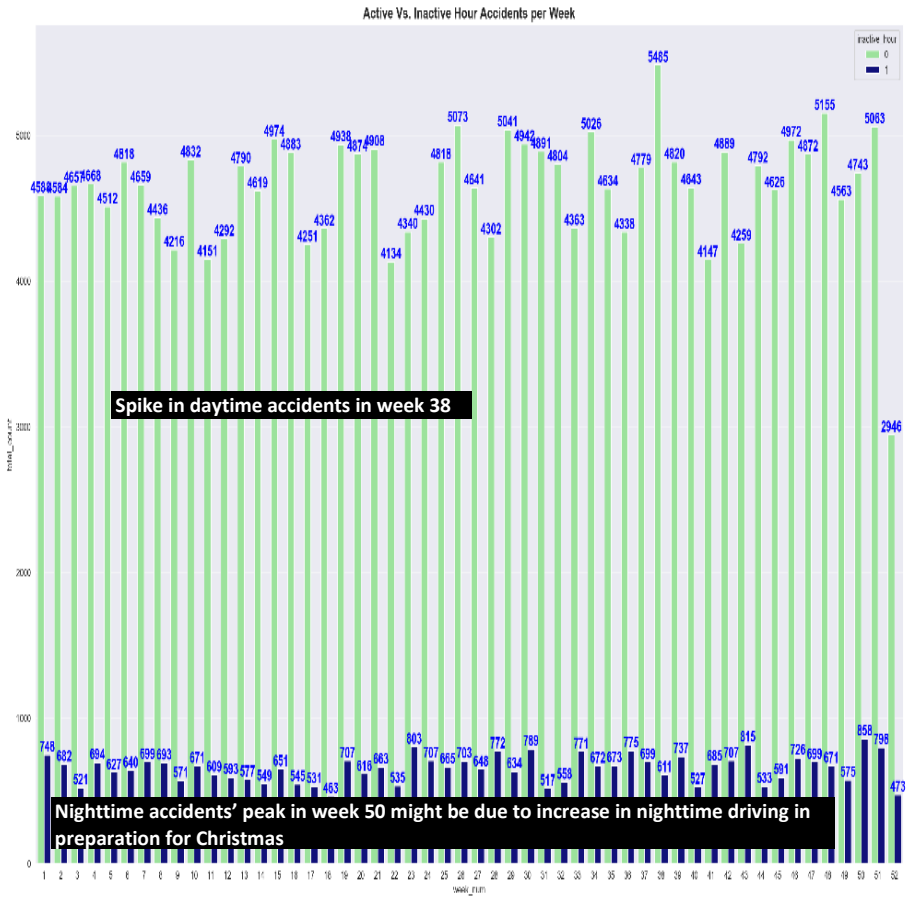


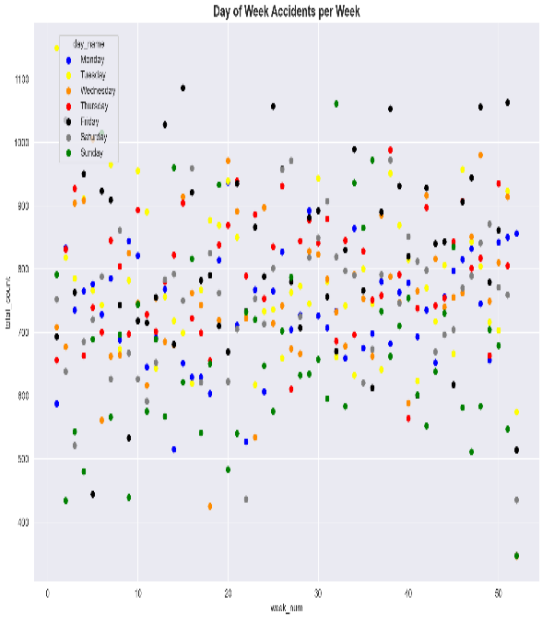
WEEKLY OBSERVATIONS



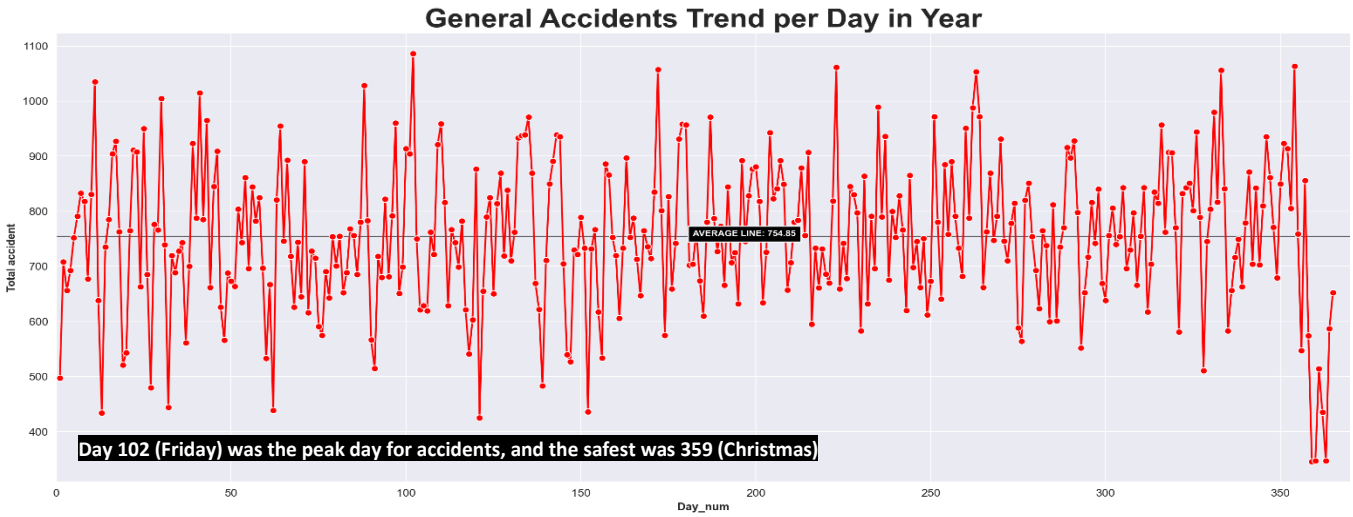

Weekend Vs Weekday Accidents per Week
is_weekend
0
1
week_num
Week 52 was the safest for weekdays and weekends probably because people spent more time with family and less time on the road at that time of the year
Weekend accidents was highest in week 32 (the first week of a new league season)

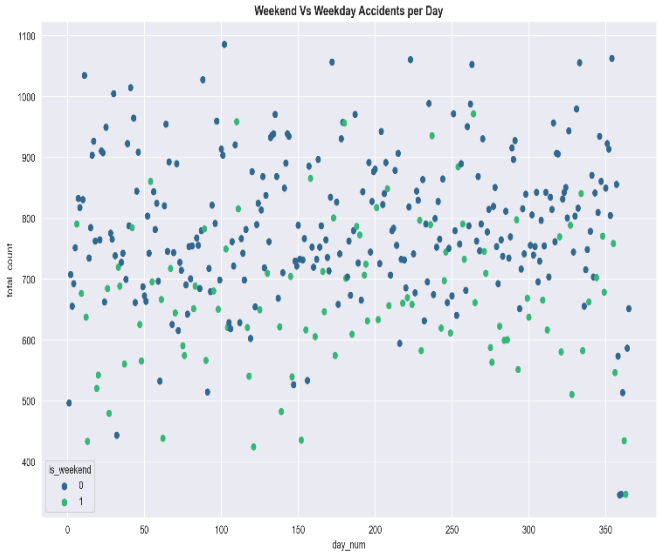

Seasonal Accidents per Week
season
winter
spring
summer
autumn
week_num

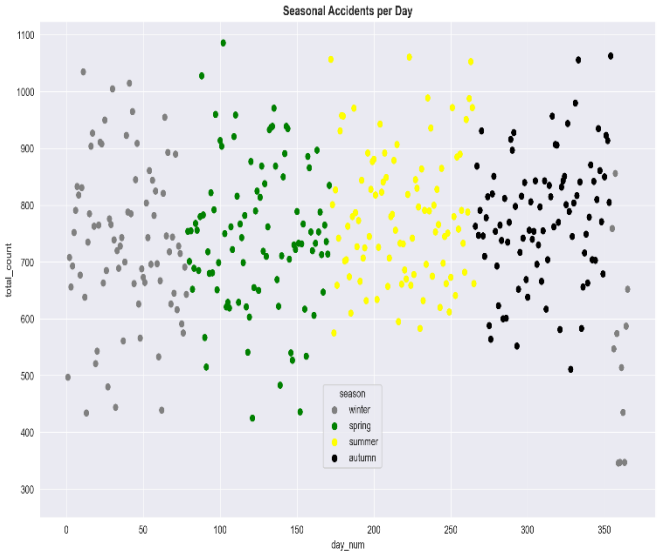

Active Vs. Inactive Hour Accidents per Week
inactive_hour
0
1
week_num
Spike in daytime accidents in week 38
Nighttime accidents' peak in week 50 might be due to increase in nighttime driving in preparation for Christmas

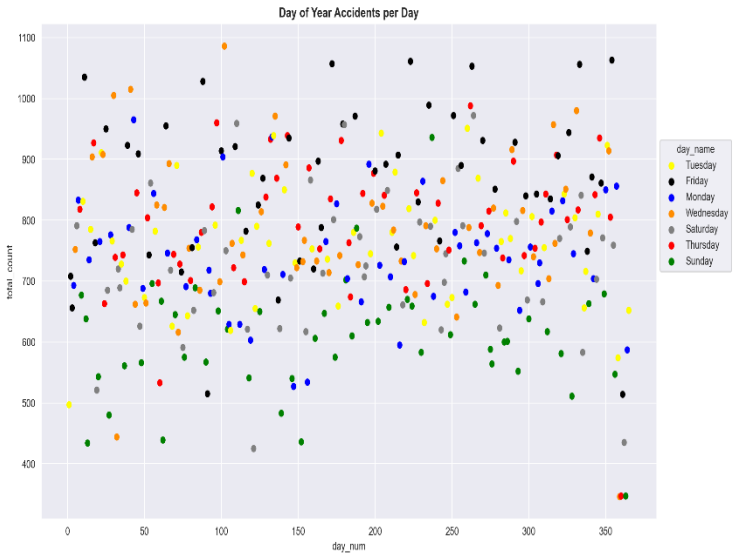

Day of Week Accidents per Week
day_name
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
week_num

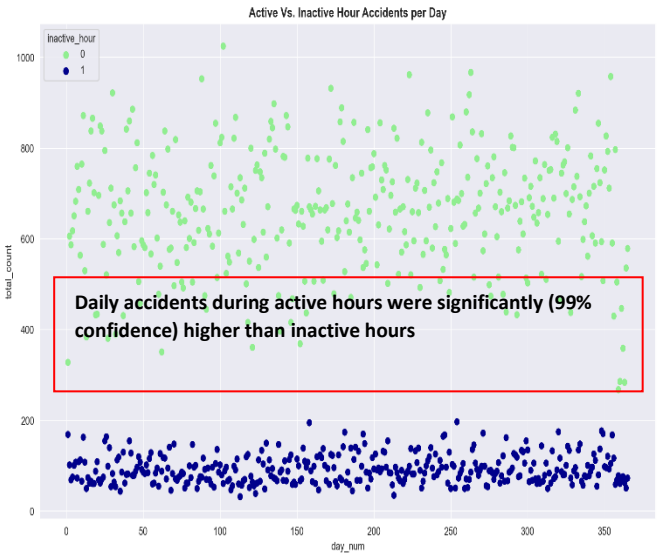
DAILY OBSERVATIONS



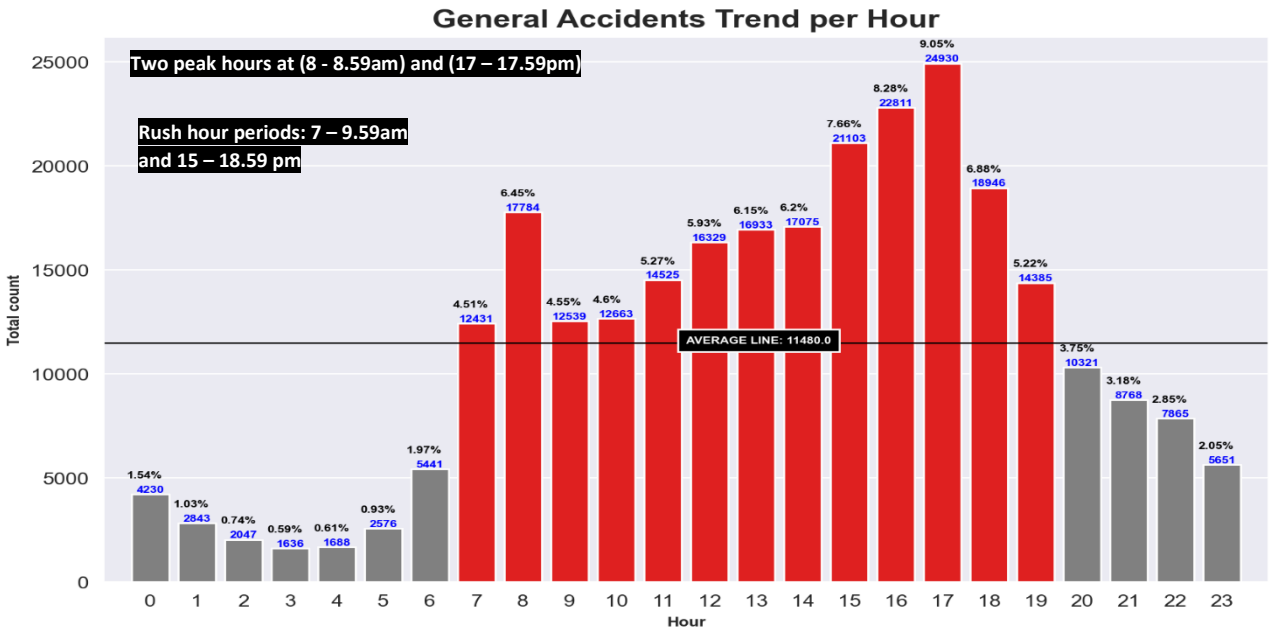


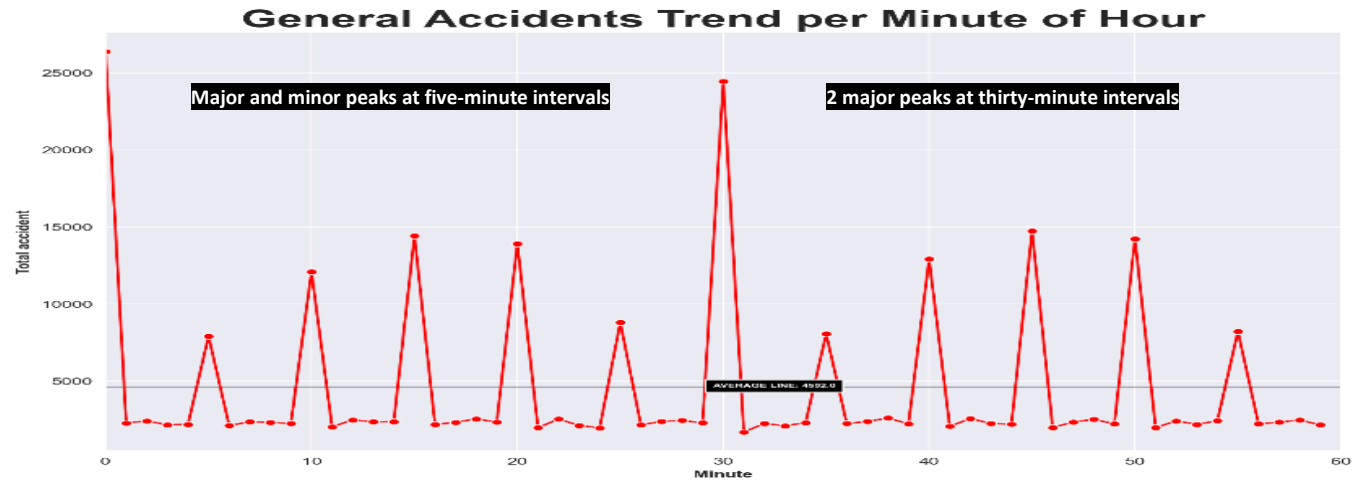


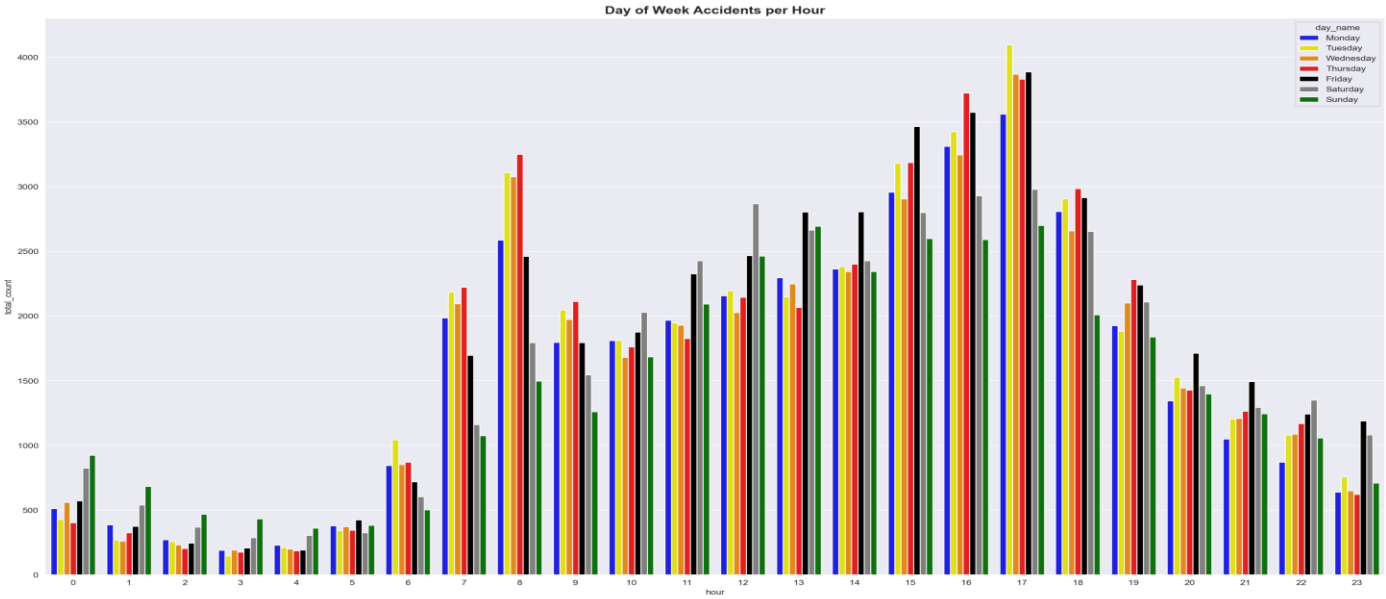




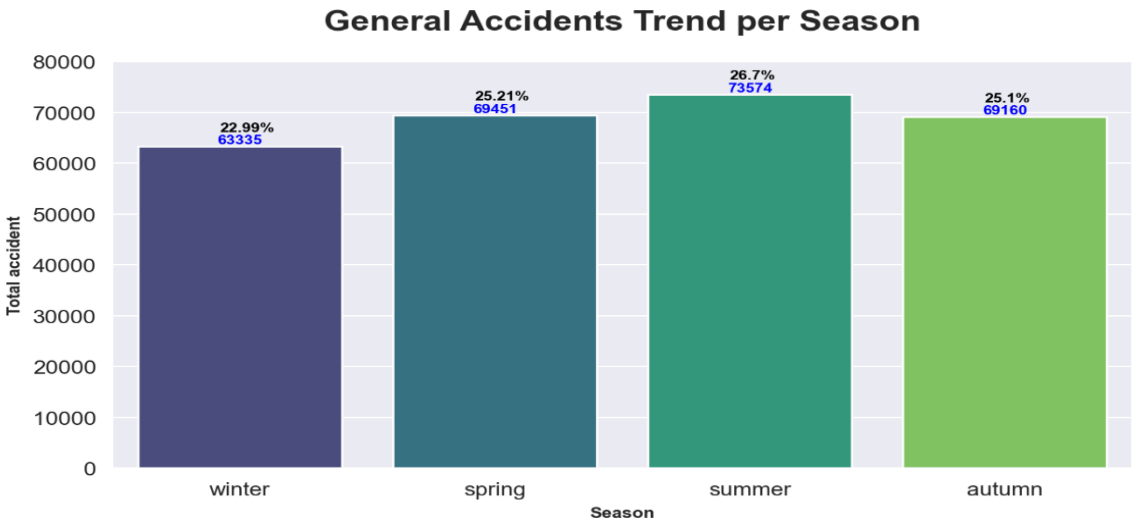
HOURLY OBSERVATIONS

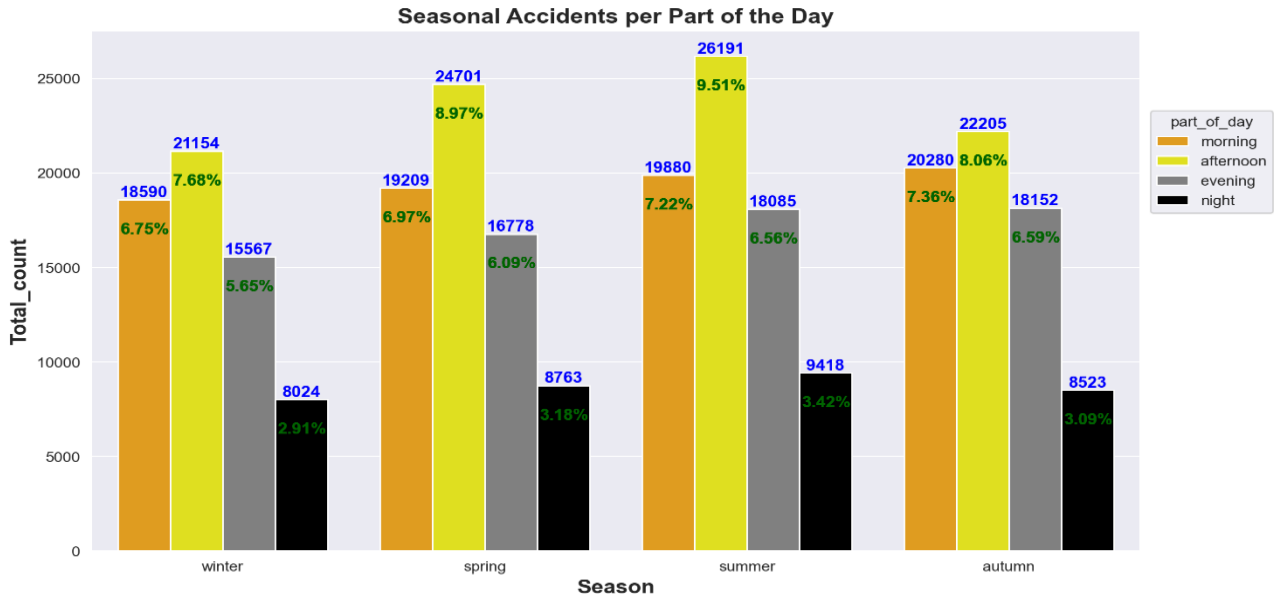


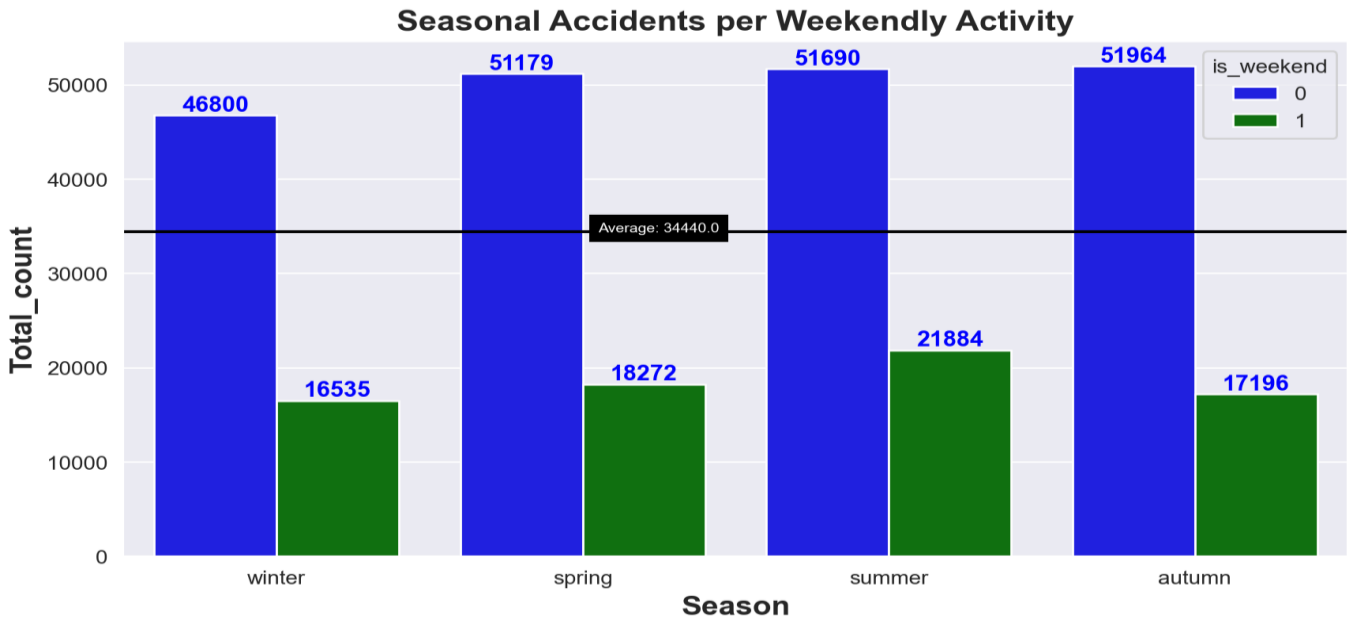


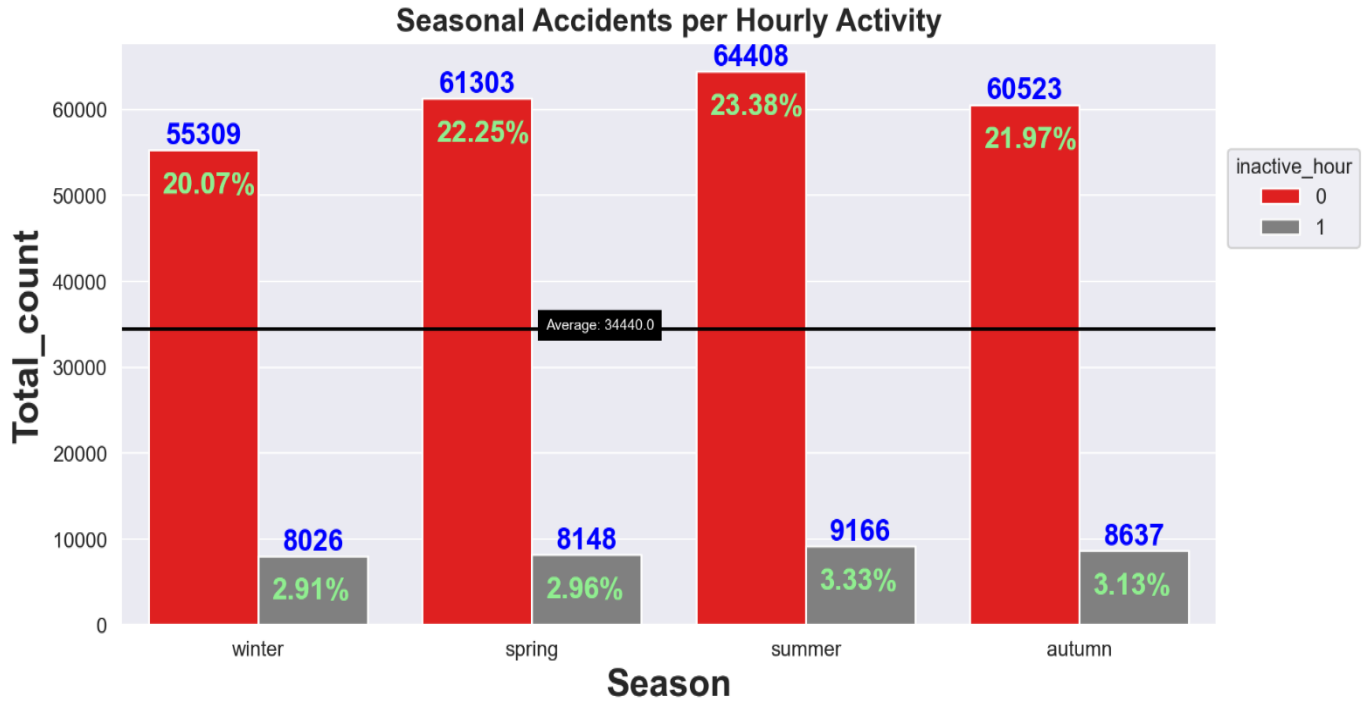


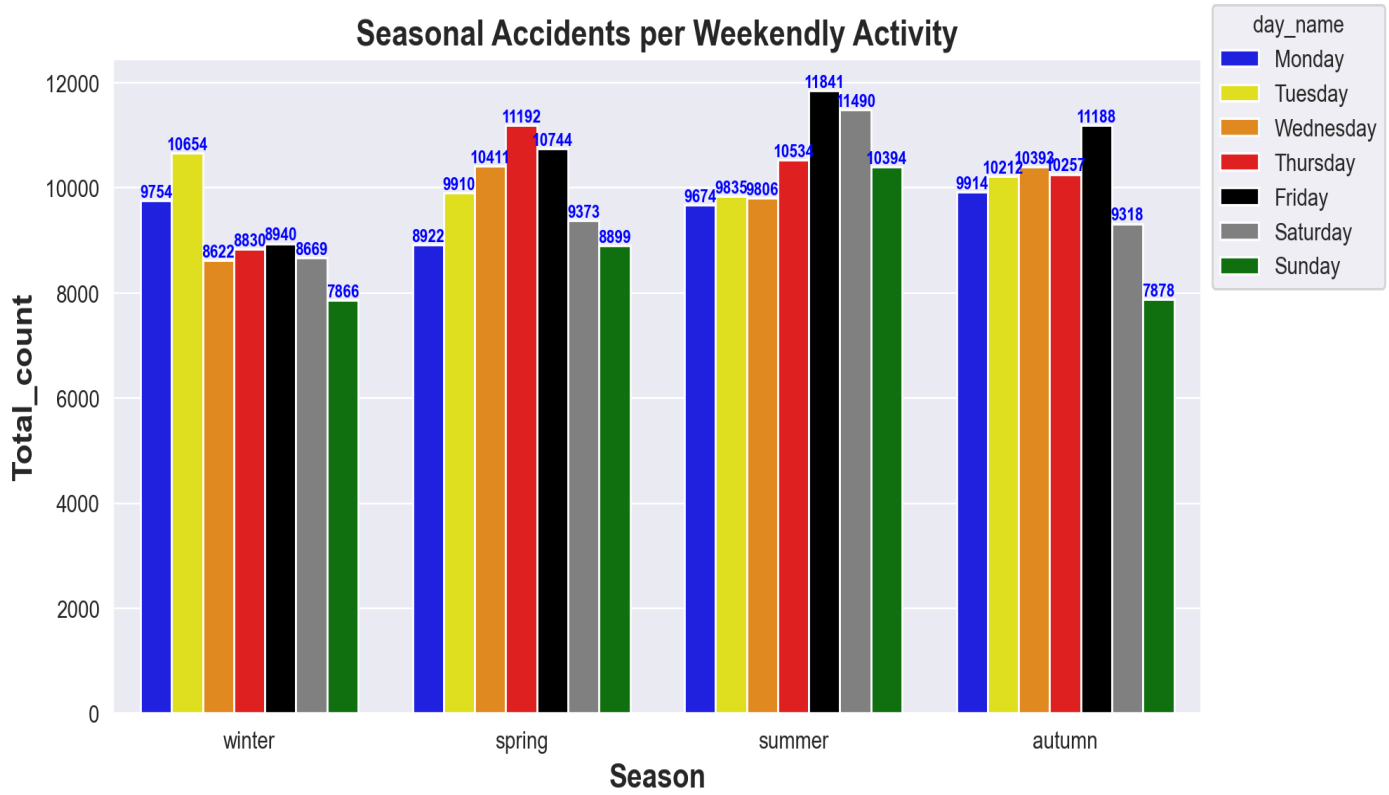
SEASONAL OBSERVATIONS



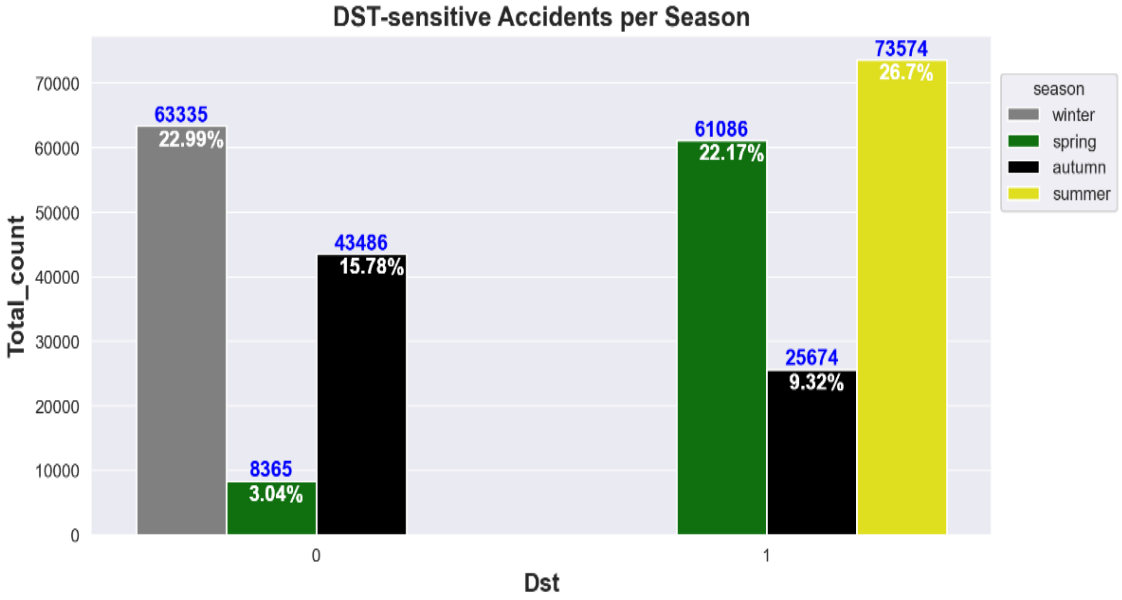


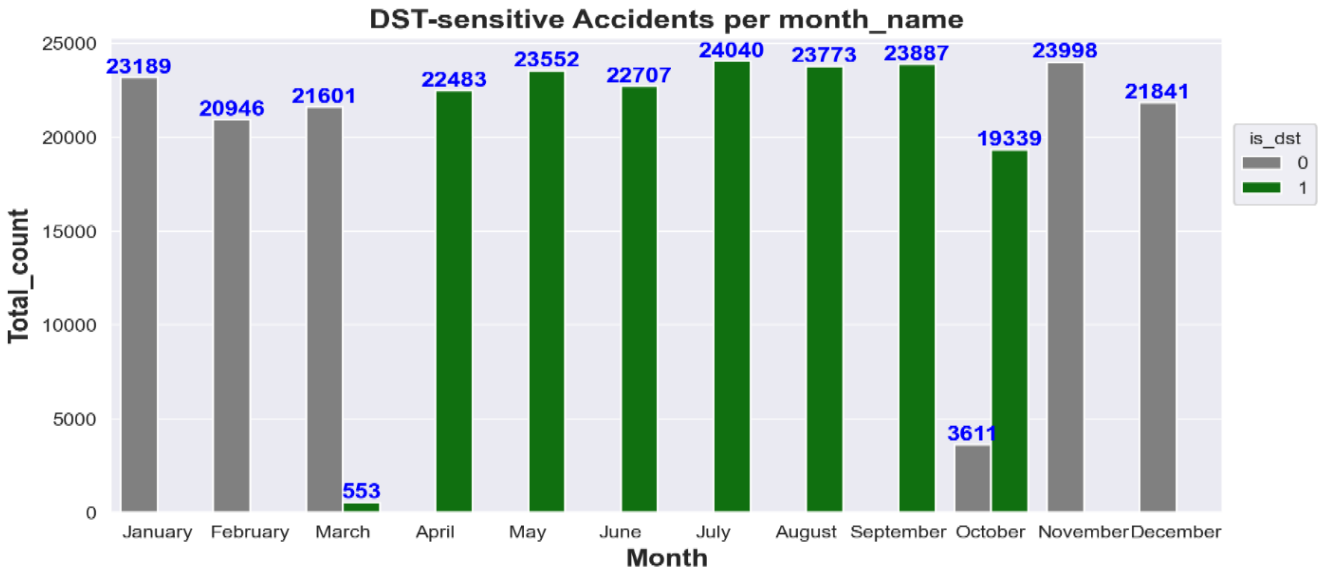


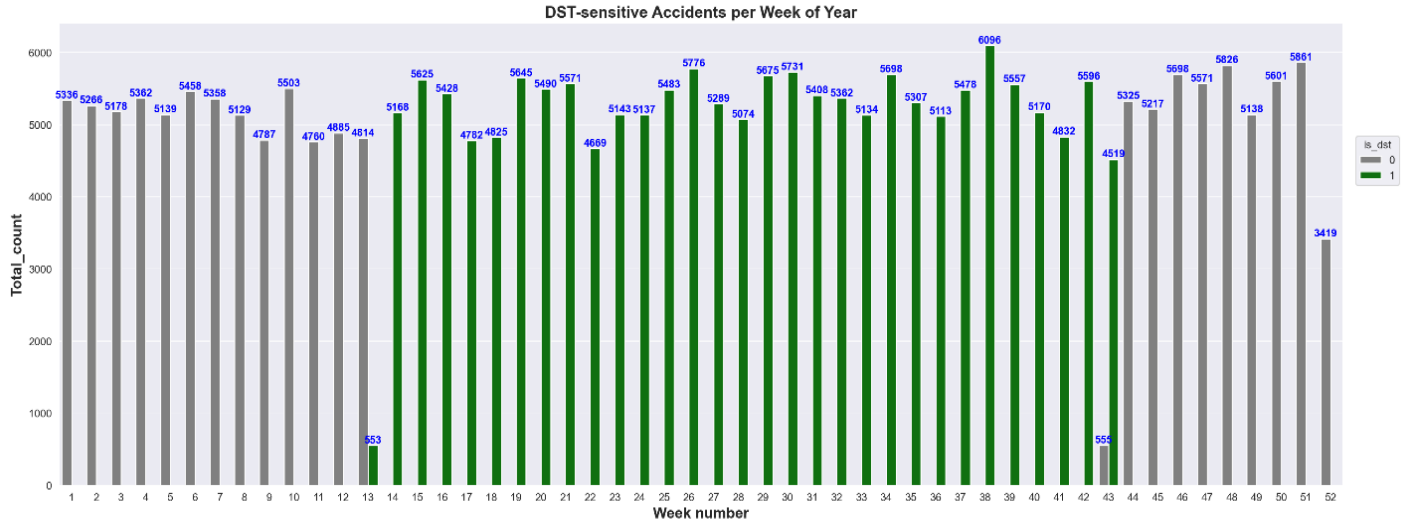

Seasonal Accidents per Hourly Activity
Total_count
Season
winter
spring
summer
autumn
55309
20.07%
8026
2.91%
61303
22.25%
8148
2.96%
64408
23.38%
9166
3.33%
60523
21.97%
8637
3.13%
Average: 34440.0
inactive_hour
0
1

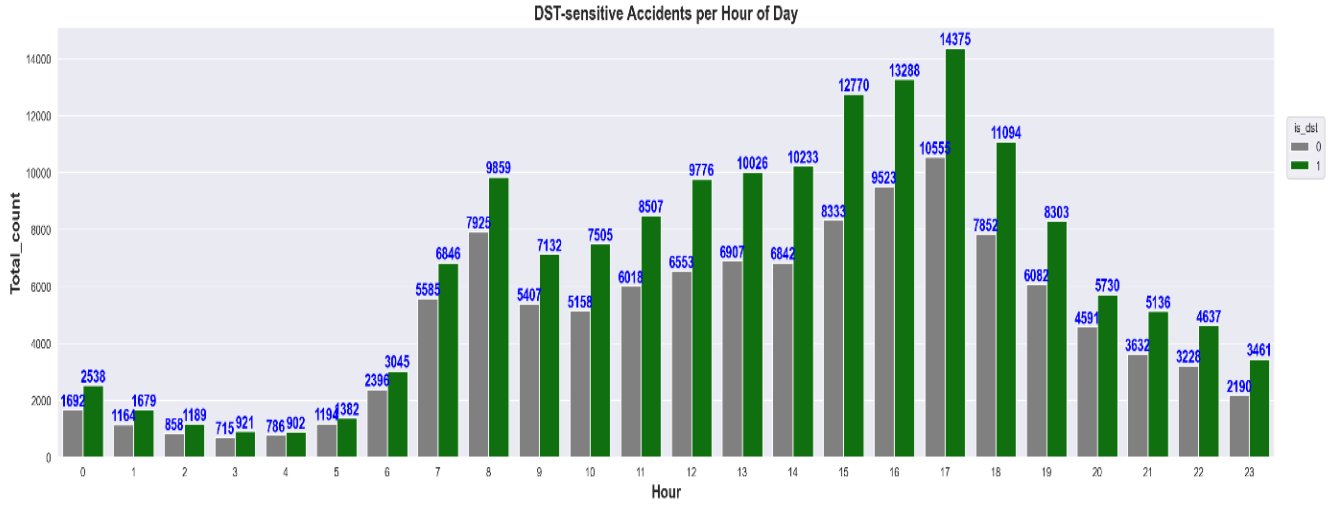

Seasonal Accidents per Weekendly Activity
Total_count
Season
winter
spring
summer
autumn
day_name
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
9754
10654
8622
8830
8940
8669
7866
8922
9910
10411
11192
10744
9373
8899
9674
9835
9806
10534
11841
11490
10394
9914
10212
10393
10257
11188
9318
7878

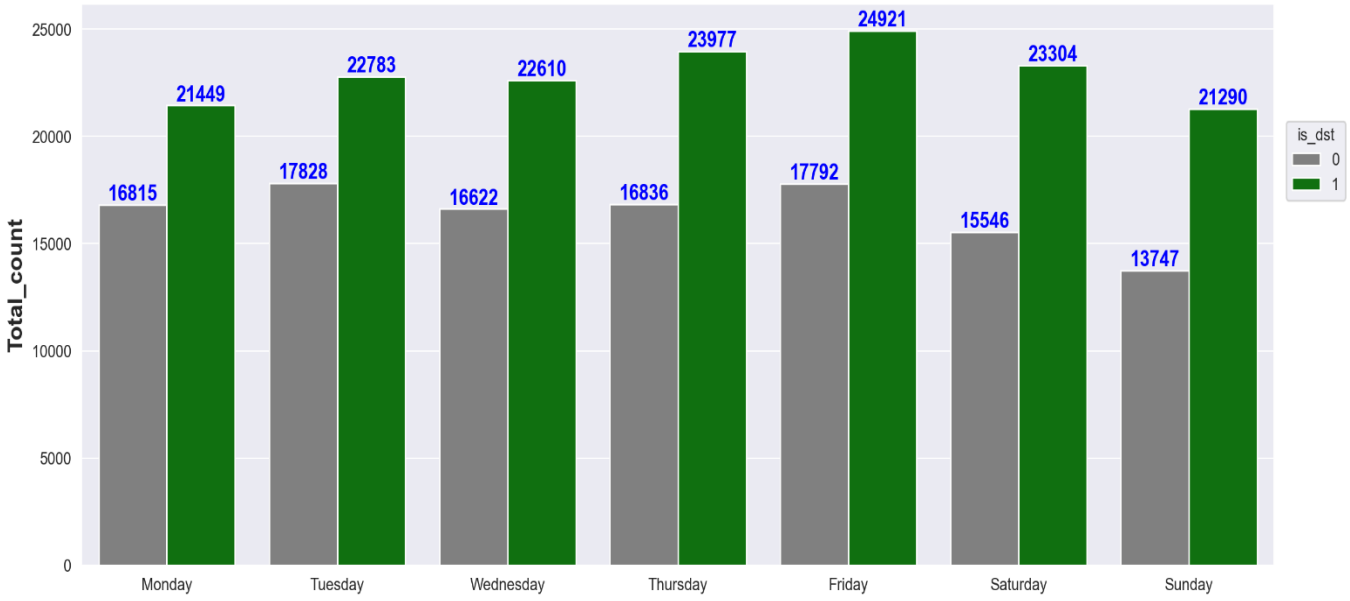
DAYLIGHT SAVING TIME (DST) OBSERVATIONS

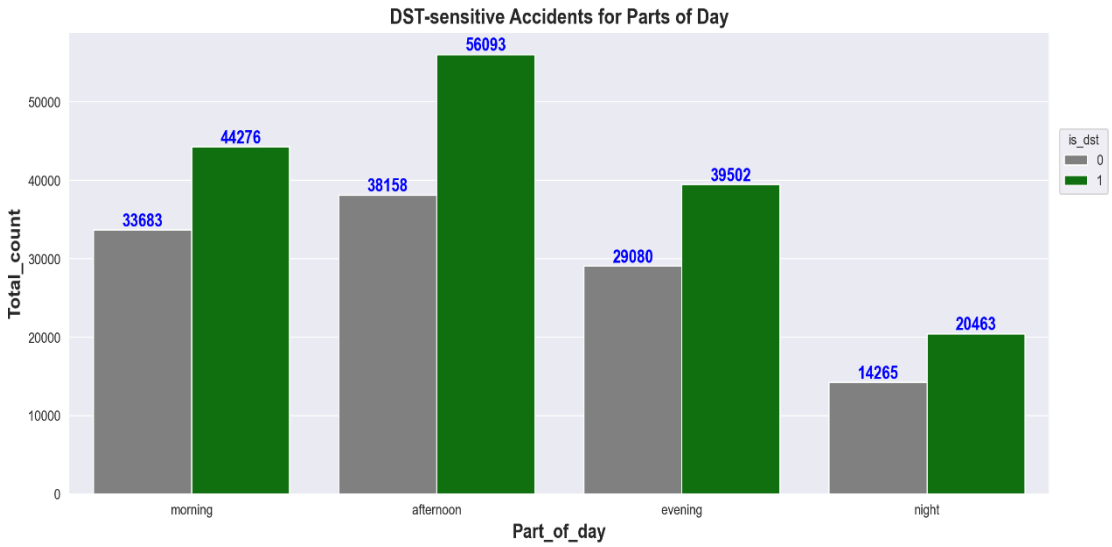







DST-sensitive Accidents per Hour of Day
Total_count
Hour
is_dst
0
1
1692
2538
1164
1679
858
1189
715
921
786
902
1194
1382
2396
3045
5585
6846
7925
9859
5407
7132
5158
7505
6018
8507
6553
9776
6907
10026
6842
10233
8333
12770
9523
13288
10555
14375
7852
11094
6082
8303
4591
5730
3632
5136
3228
4637
2190
3461

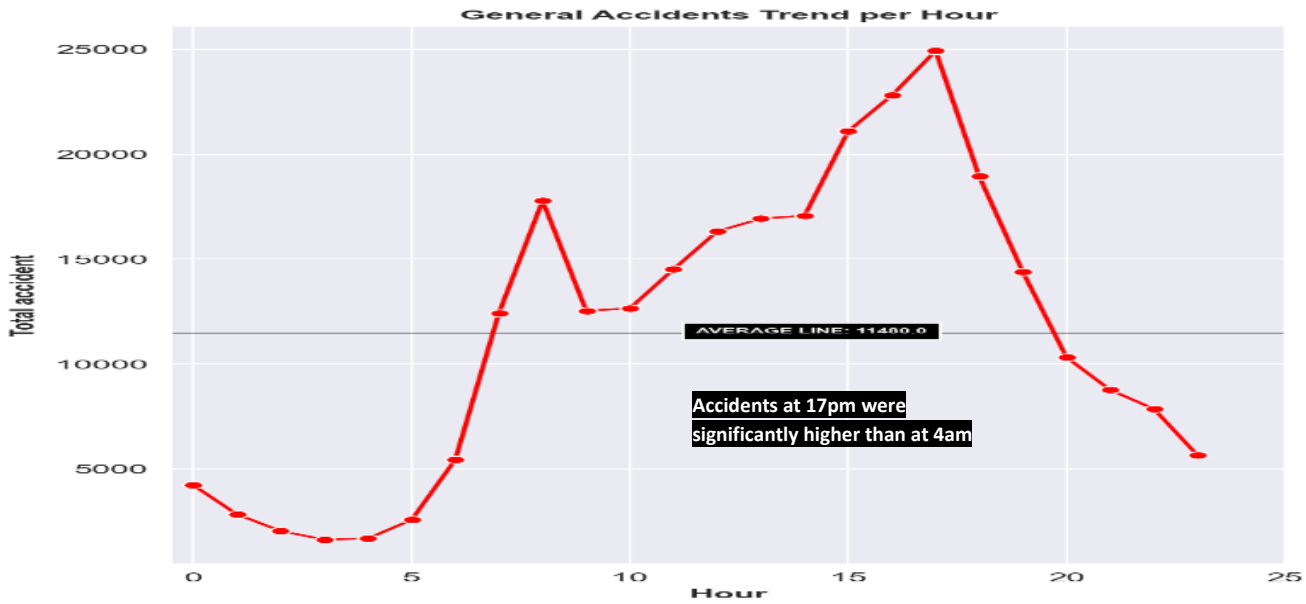

Total_count
is_dst
0
1
16815
21449
17828
22783
16622
22610
16836
23977
17792
24921
15546
23304
13747
21290
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday


DST-sensitive Accidents for Parts of Day
Total_count
Part_of_day
is_dst
0
1
33683
44276
38158
56093
29080
39502
14265
20463
morning
afternoon
evening
night

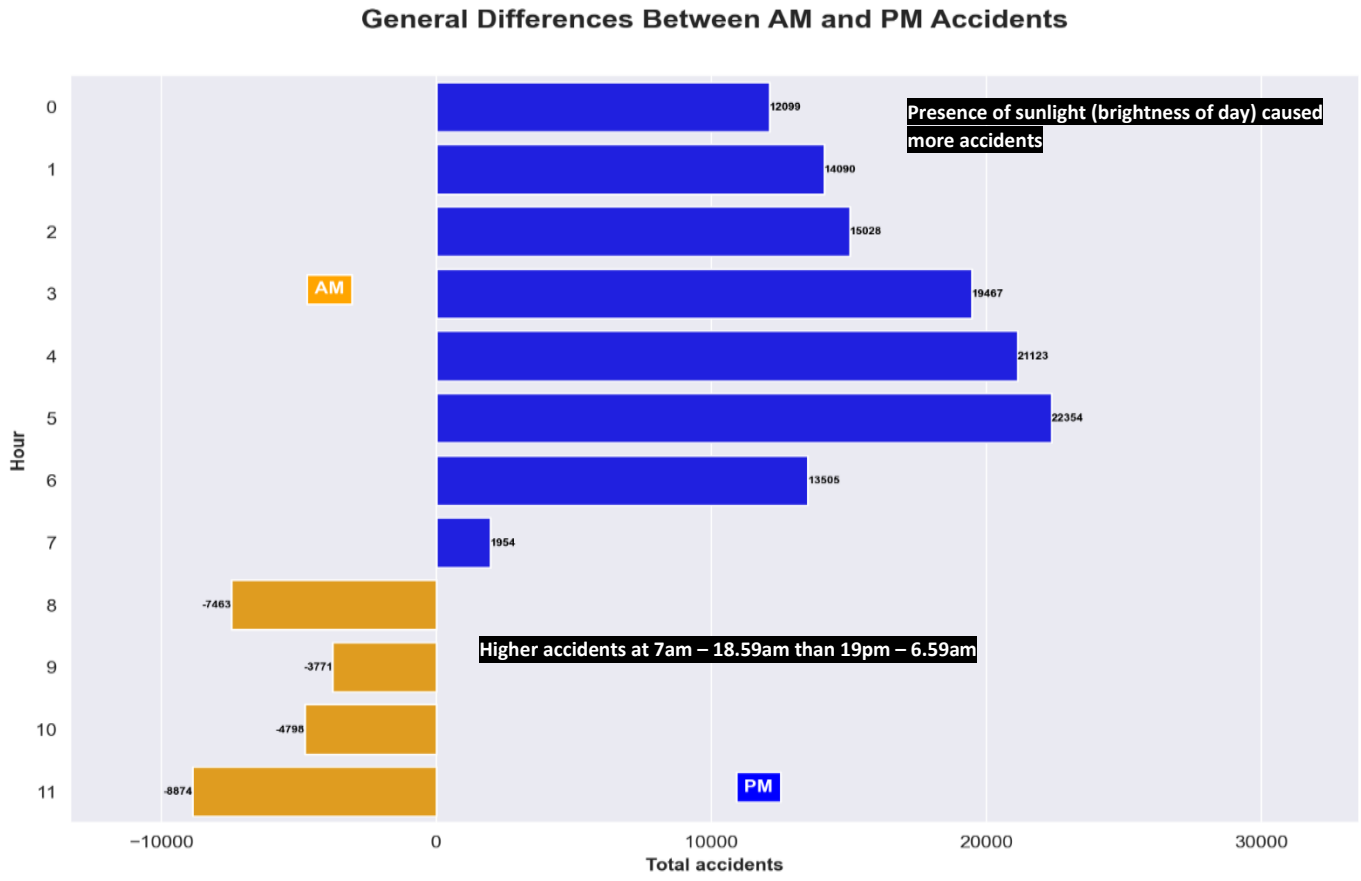
INFERENTIAL STATISTICS

Statistical testing was performed using a pooled variance, and significance level was established when the mean_difference >= 2Xpooled_std (95% confidence), or 3Xpooled_std (99% confidence).

TOTAL ACCIDENTS:



a. SIGNIFICANT HOURS OF DAY



H0: Monthly and hourly accidents between 7am and 18.59pm each were not higher than between 19pm and 6.59am.

At 95% confidence, there is evidence to show that H0 is false. Hence, we reject H0 and accept that monthly and hourly accidents between 7am and 18.59pm each were significantly higher than between 19pm and 6.59am.

H0: Afternoon daily and hourly accidents each were not higher than nighttime accidents.

There is evidence to show that H0 is false. Therefore, we reject H0 and accept that afternoon daily (95% confidence) and hourly (99% confidence) accidents were significantly higher than evening and nighttime accidents each.

H0: active time hourly (7am – 21.59pm) daily accidents were not higher than inactive hourly (22pm – 6.59am).

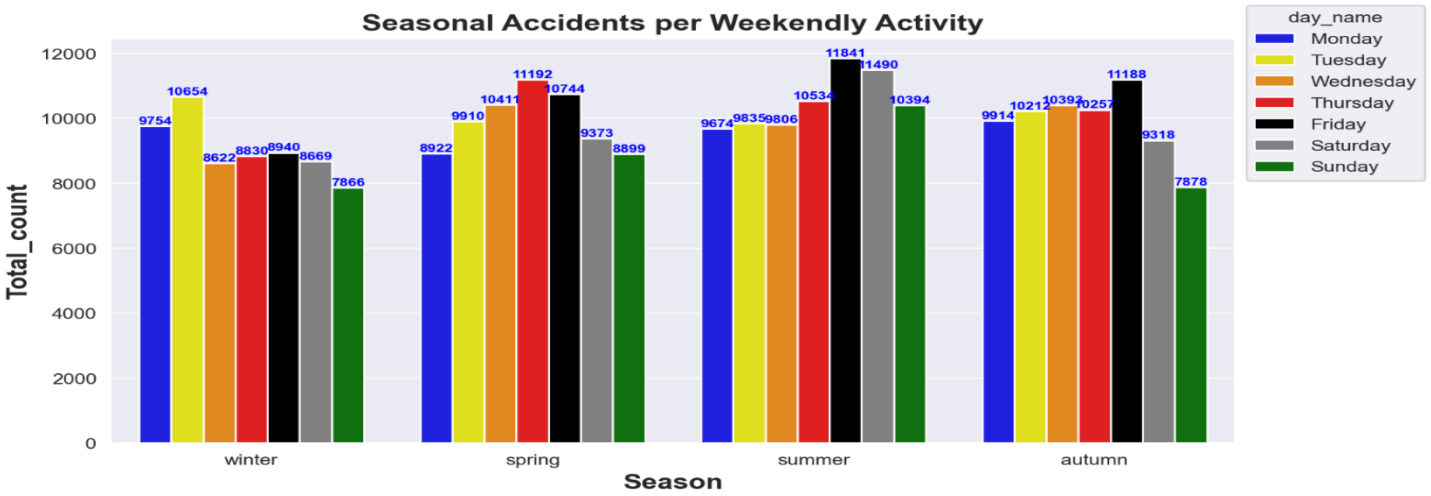
At 99% confidence, there is evidence to show that H0 is false. Hence, we reject H0 and accept that active time hourly accidents were significantly higher than between 19pm and 6.59am.

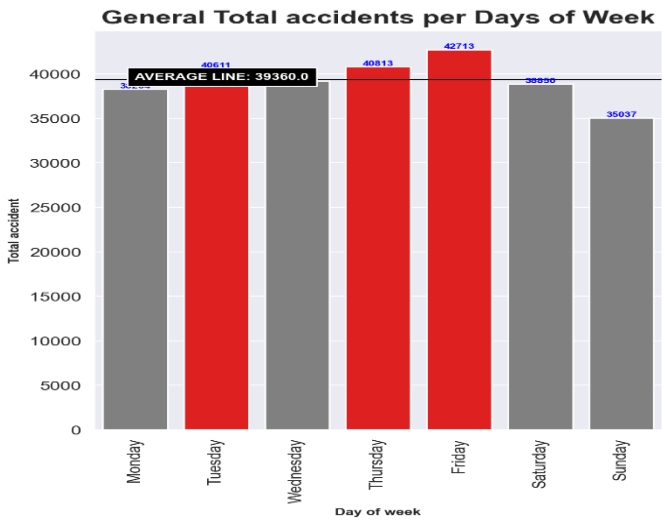
H0: Monthly afternoon accidents were not higher than evening accidents.

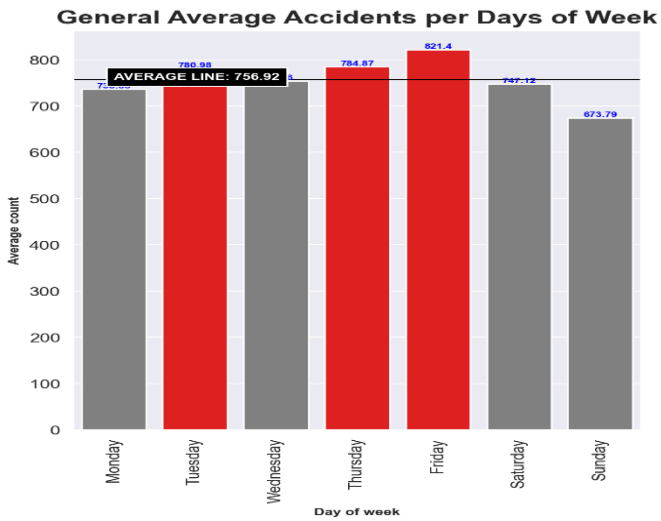
At 99% confidence, there is evidence to show that H0 is false. Therefore, we reject H0 and accept that monthly afternoon accidents were significantly higher than evening accidents.

Interpretation: These results suggest that activeness of the part of day increased the likelihood of an accident compared to inactivity associated with nighttime.

b. <u>SIGNIFICANT DAYS OF WEEK</u>







H0: monthly and weekly weekday accidents were not higher than weekends

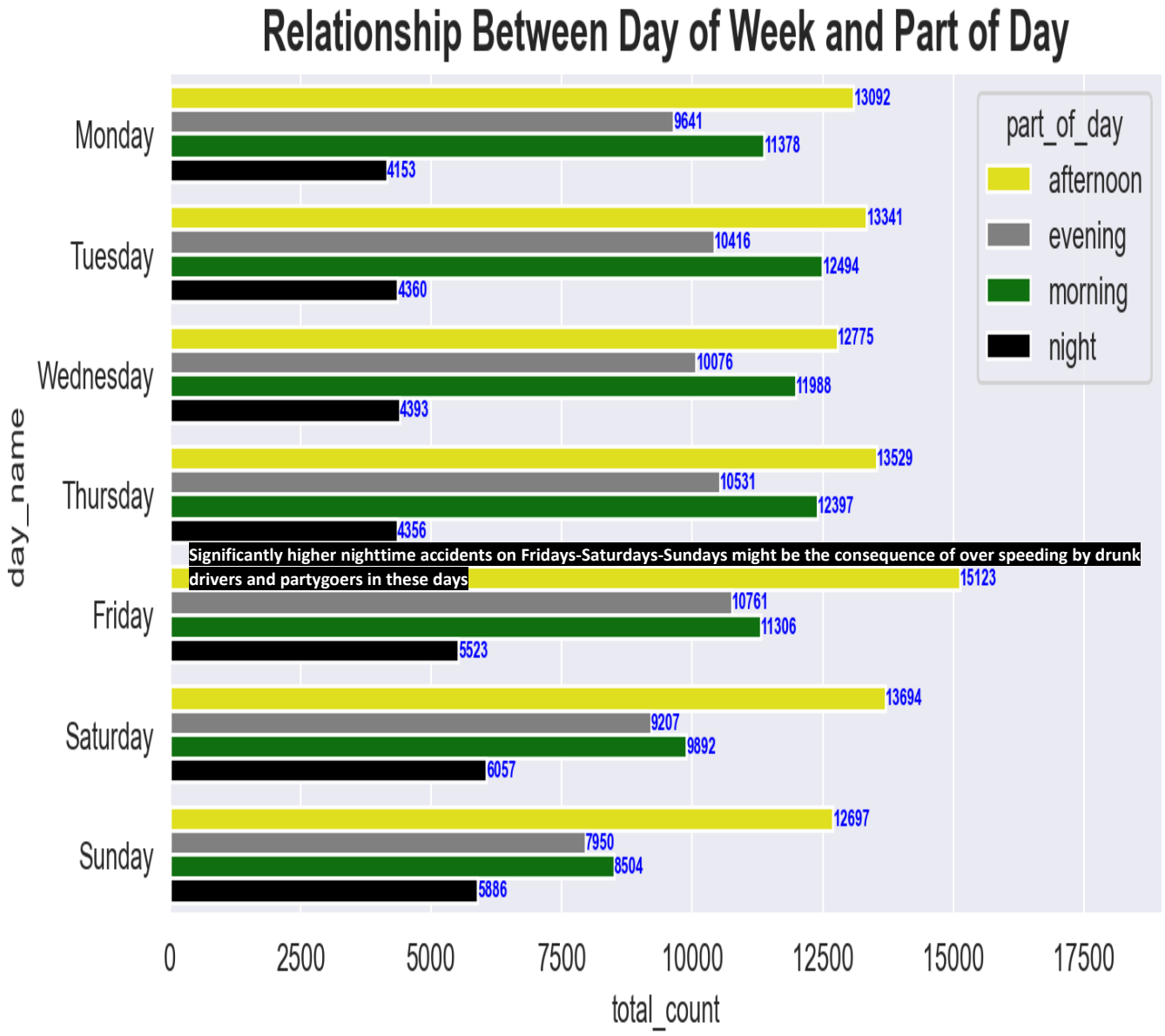
At 99% confidence, there is evidence to show that H0 is false. Hence, we reject H0 and accept that monthly and weekly weekday accidents were higher than monthly weekend accidents.

Interpretation: These results suggest that overall accidents were more likely to occur on a weekday than a weekend probably because more people commute on weekdays (e.g to schools, banks).

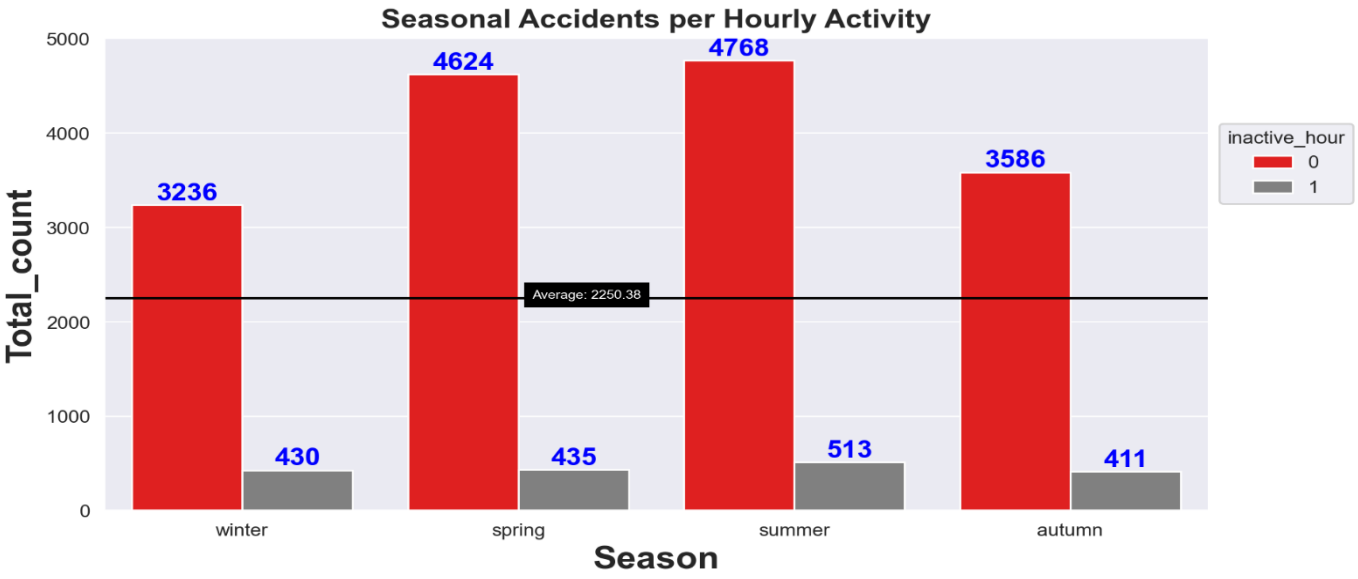
H0: nighttime accidents were not higher on Fridays-Saturdays-Sundays than Mondays-Tuesdays-Wednesdays-Thursdays.

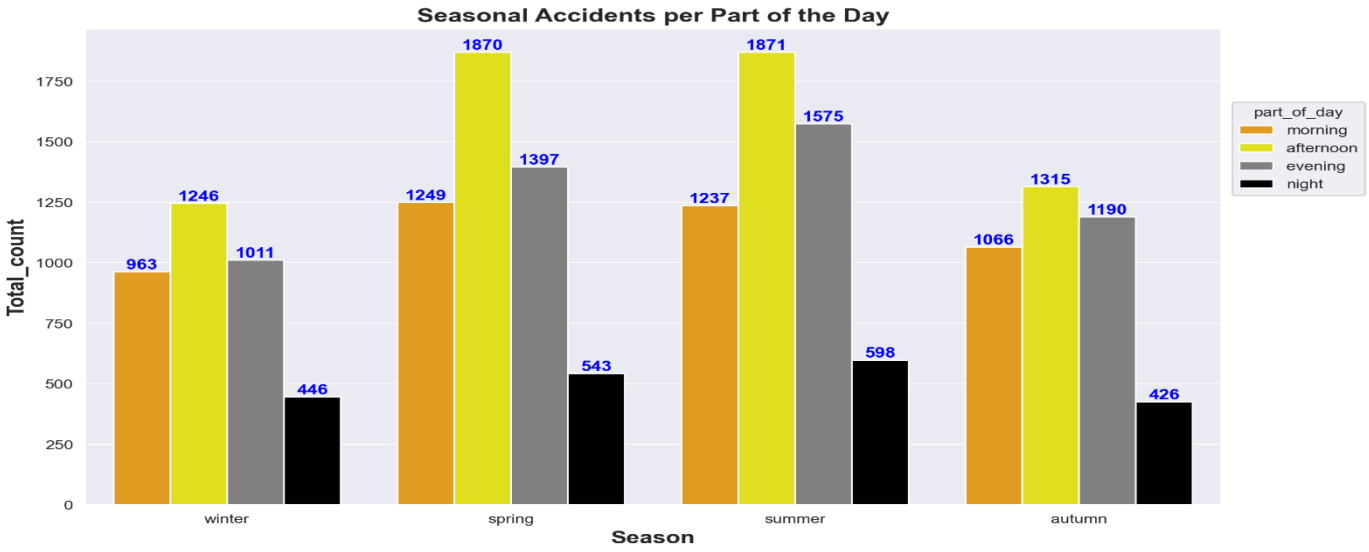
At 99% confidence, there is evidence to show that H0 is false. Hence, we reject H0 and accept that nighttime accidents on Fridays-Saturdays-Sundays were significantly higher than Mondays-Tuesdays-Wednesdays-Thursdays.

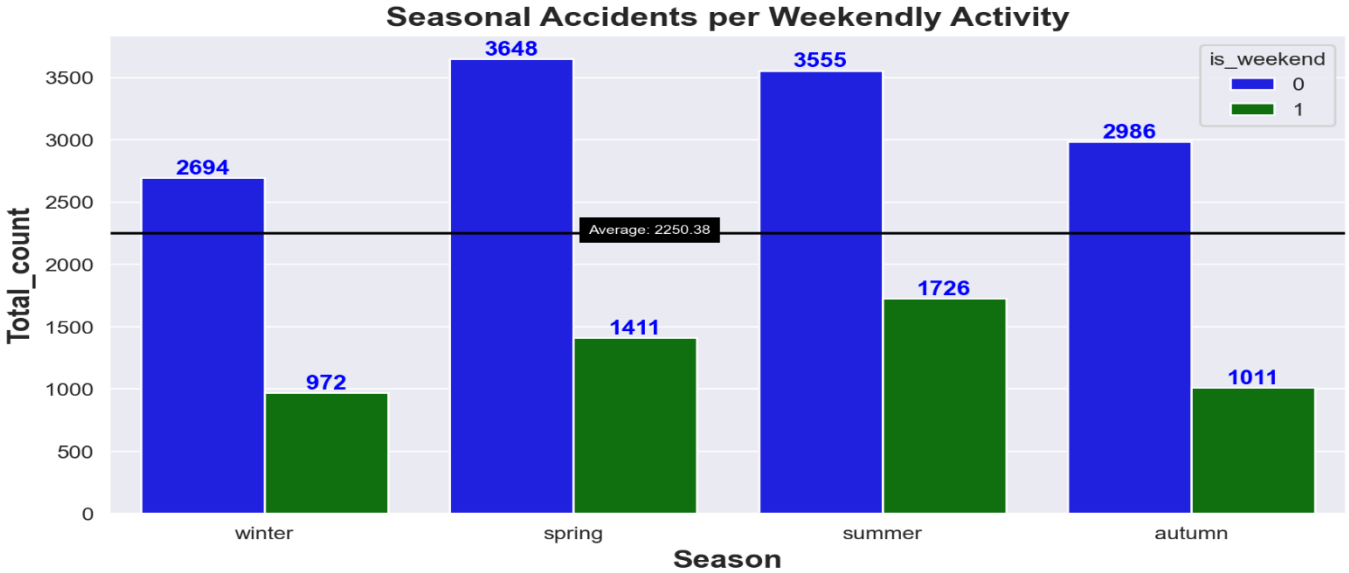
Interpretation: These results suggest that accidents at night were more likely to occur on a Friday-Saturday-Sunday than a Monday-Tuesday-Wednesday-Thursday.

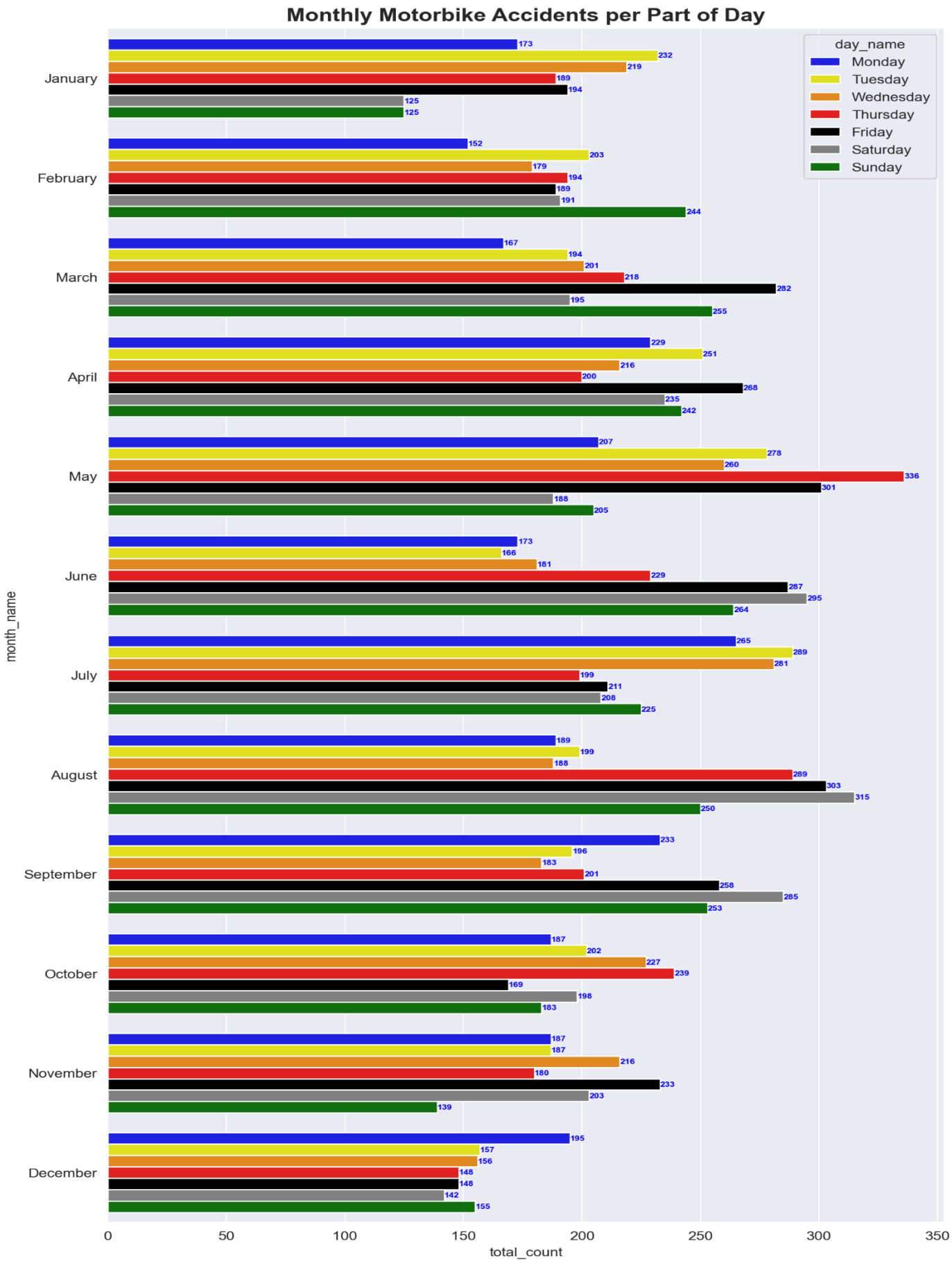


MOTORBIKE ACCIDENTS

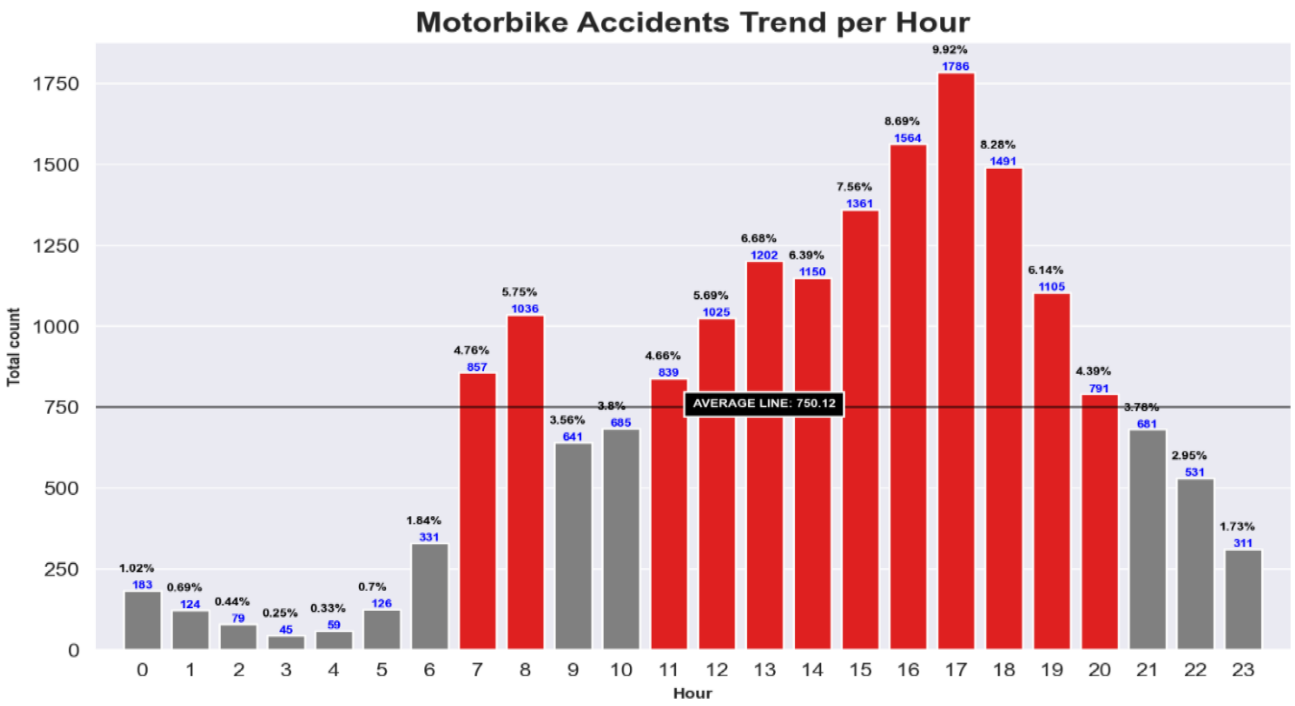


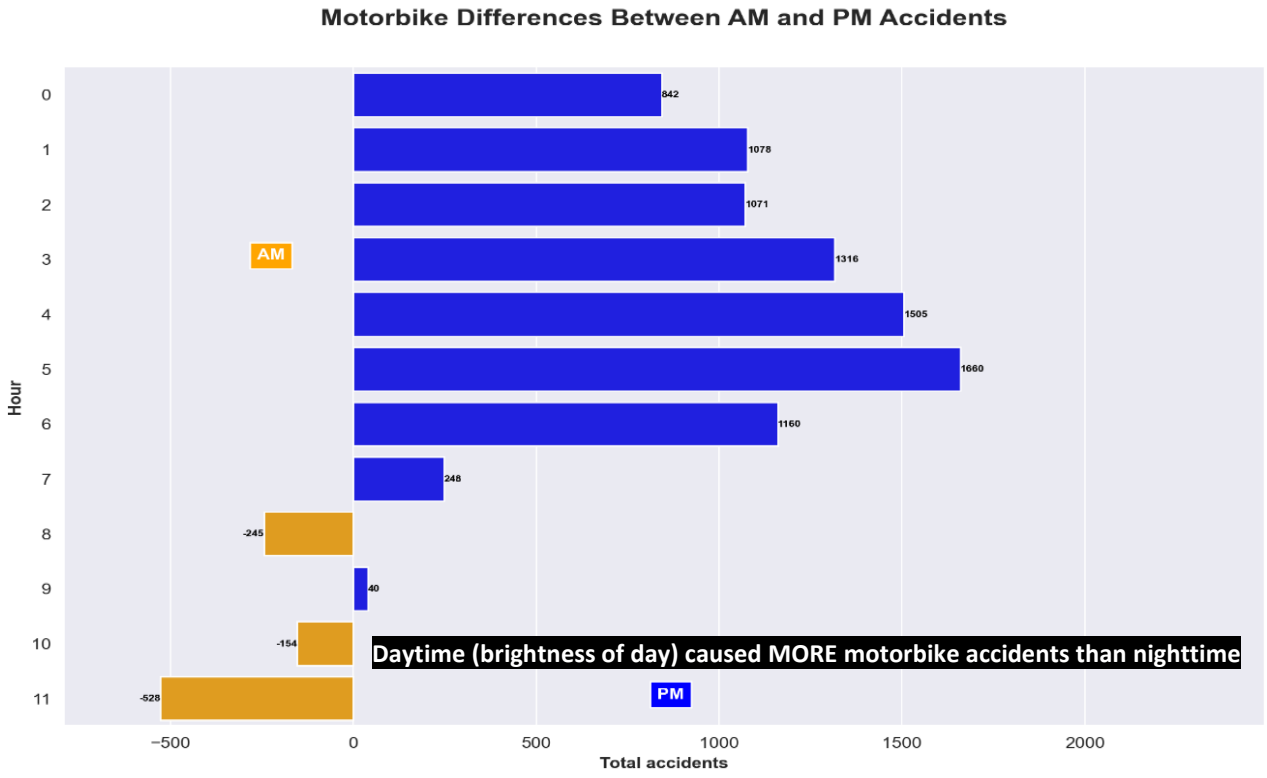





Monthly Motorbike Accidents per Part of Day
day_name
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
month_name
January
February
March
April
May
June
July
August
September
October
November
December
173
232
219
189
194
125
125
152
203
179
194
189
191
244
167
194
201
218
282
195
255
229
251
216
200
268
235
242
207
278
260
336
301
188
205
173
166
181
229
287
295
264
265
289
281
199
211
208
225
189
199
188
289
303
315
250
233
196
183
201
258
285
253
187
202
227
239
169
198
183
187
187
216
180
233
203
139
195
157
156
148
148
142
155
0
50
100
150
200
250
300
350
total_count

a. <u>SIGNIFICANT HOURS OF DAY</u>





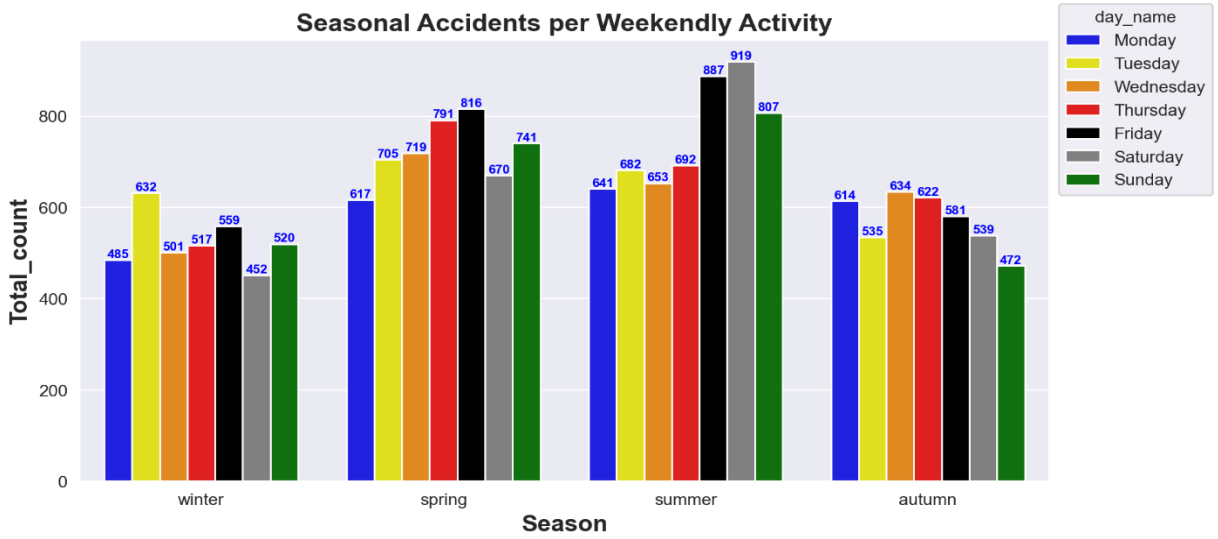
H0: total and monthly motorbike accidents between 12pm and 19.59pm each were not higher than between 0am and 7.59pm.

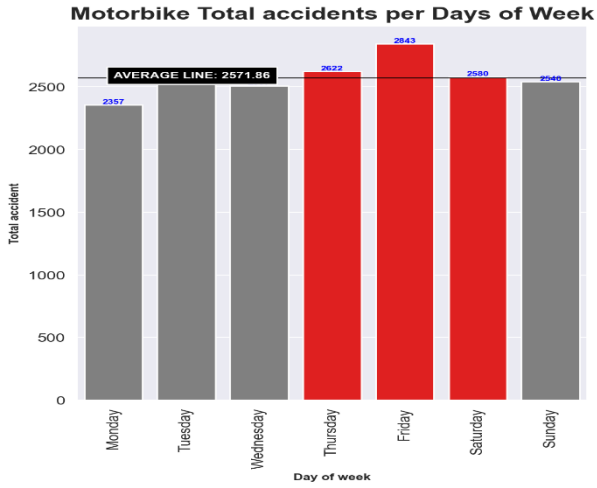
At 95% confidence, there is statistically significant evidence to show that H0 is false. We reject H0 and accept that total and monthly motorbike accidents between 12pm and 19.59pm each were significantly higher than between 0am and 7.59am.

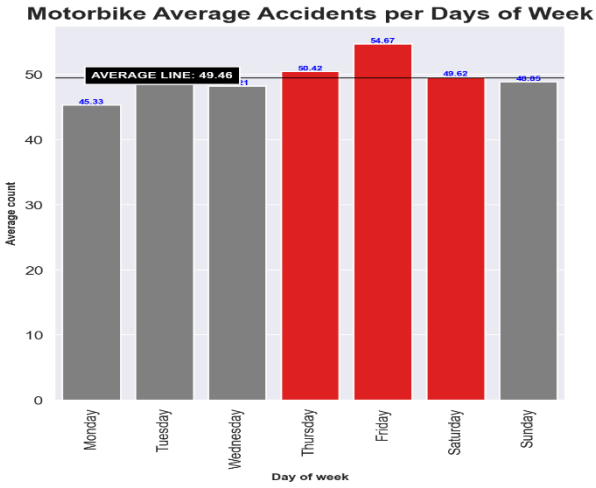
Interpretation: These results suggest that sunlight (brightness of day) which promotes human activity in general, increased the likelihood of a motorbike accident.

.

b. <u>SIGNIFICANT DAYS OF WEEK</u>





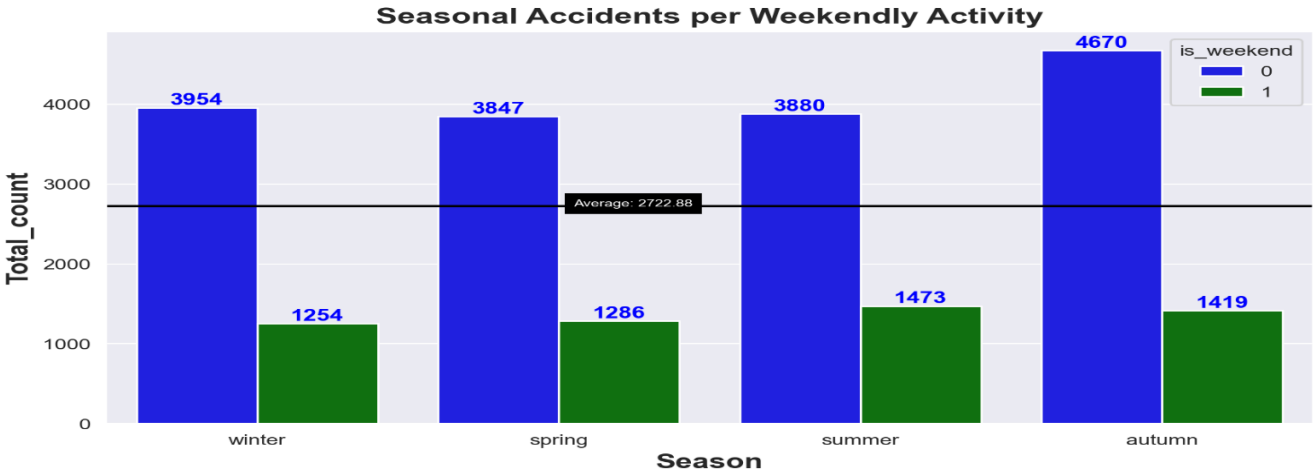


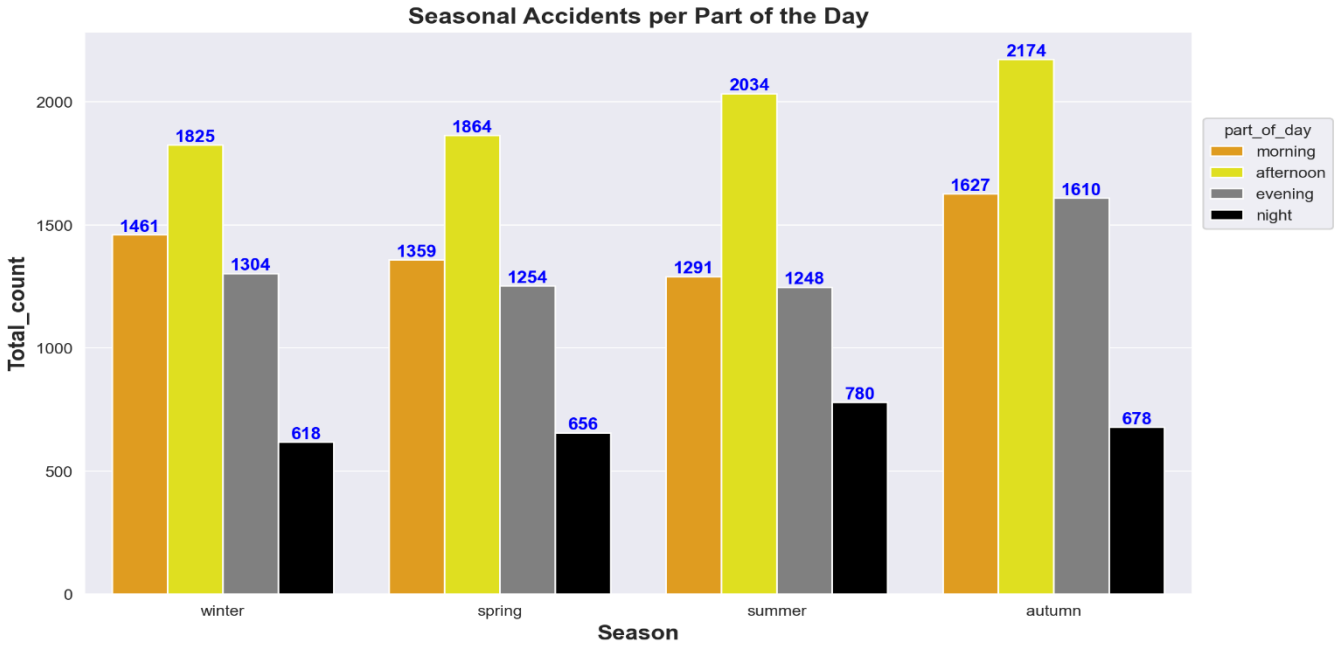
H0: Motorbike monthly and weekly accidents on weekdays each were not higher than on weekends

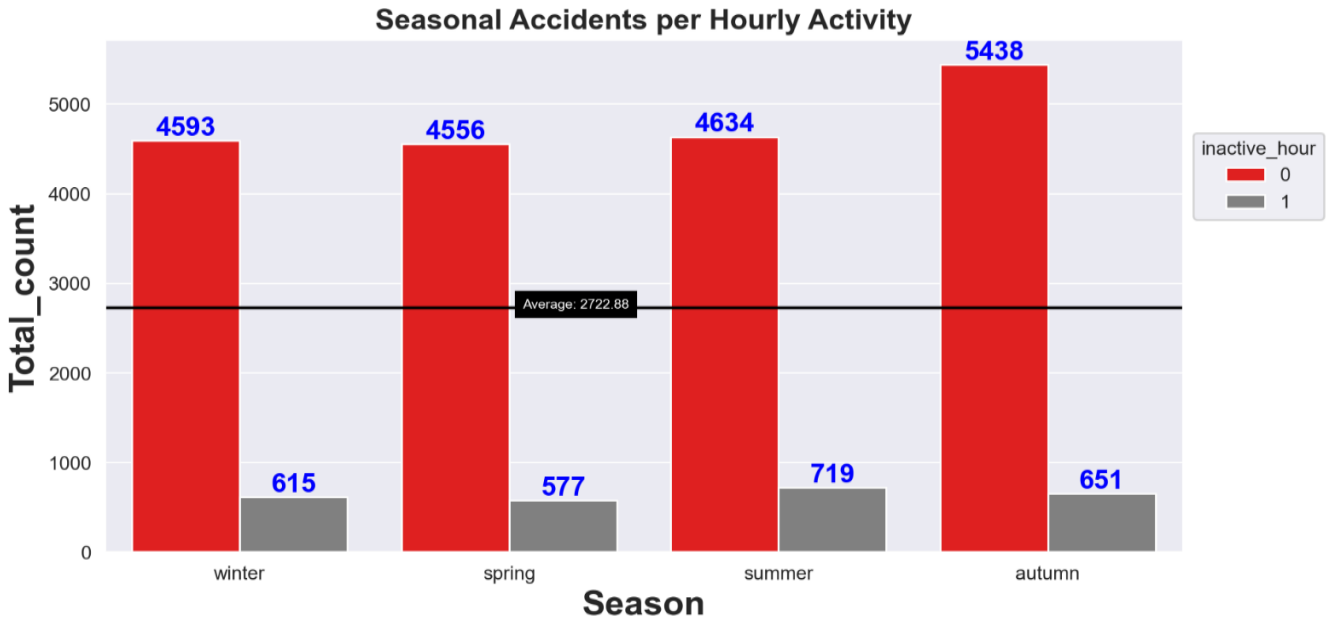
There is statistically significant evidence to show that H0 is false. Hence, we reject H0 and accept that motorbike monthly (99% confidence) and weekly (95% confidence) accidents each were higher on weekdays than weekends.

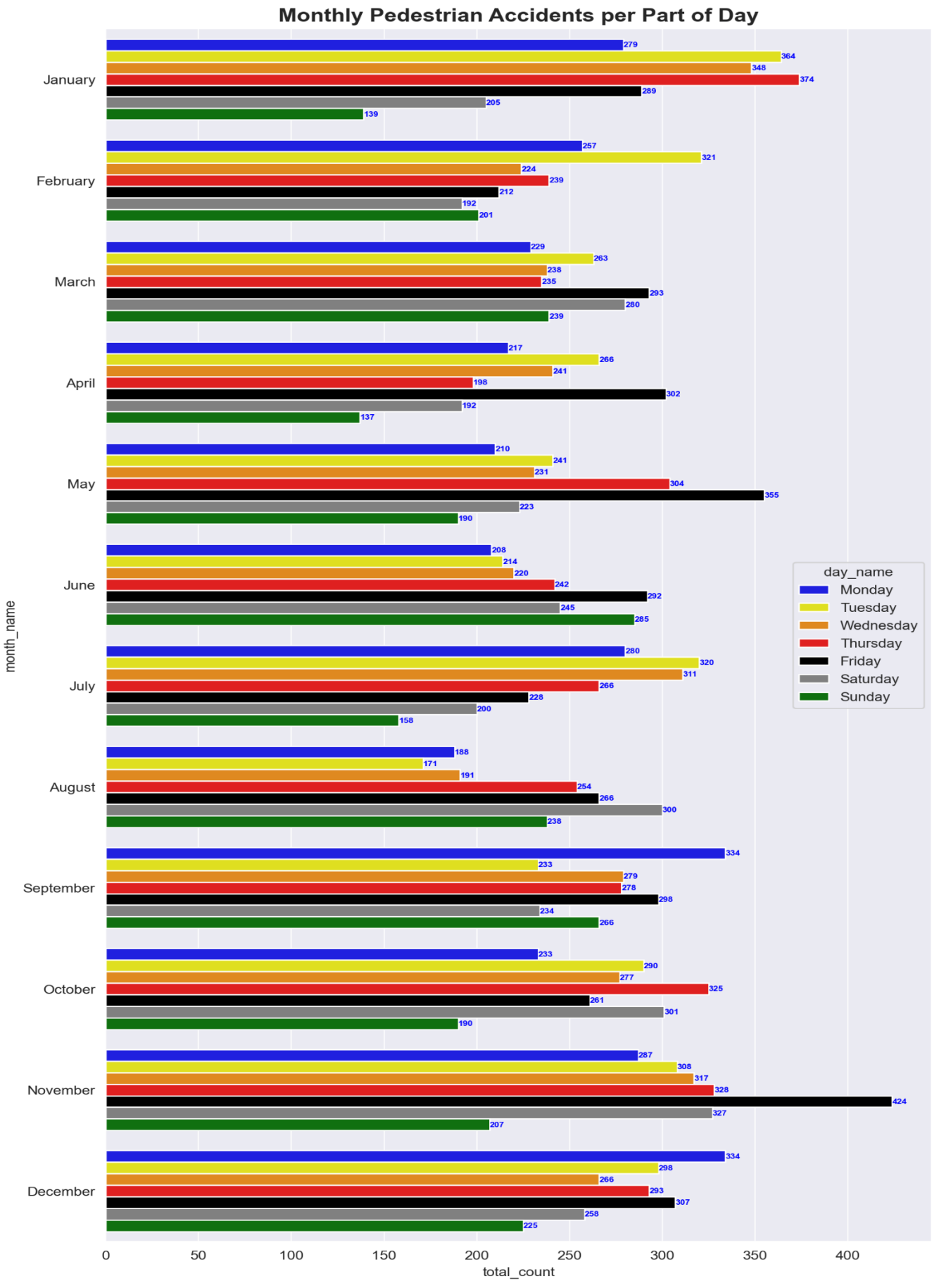
Interpretation: These results suggest that motorbike accidents were more likely on a weekday than a weekend, probably because more motorbikes are on the road on weekdays (e.g., to schools and banks) than weekends.

PEDESTRIAN ACCIDENTS

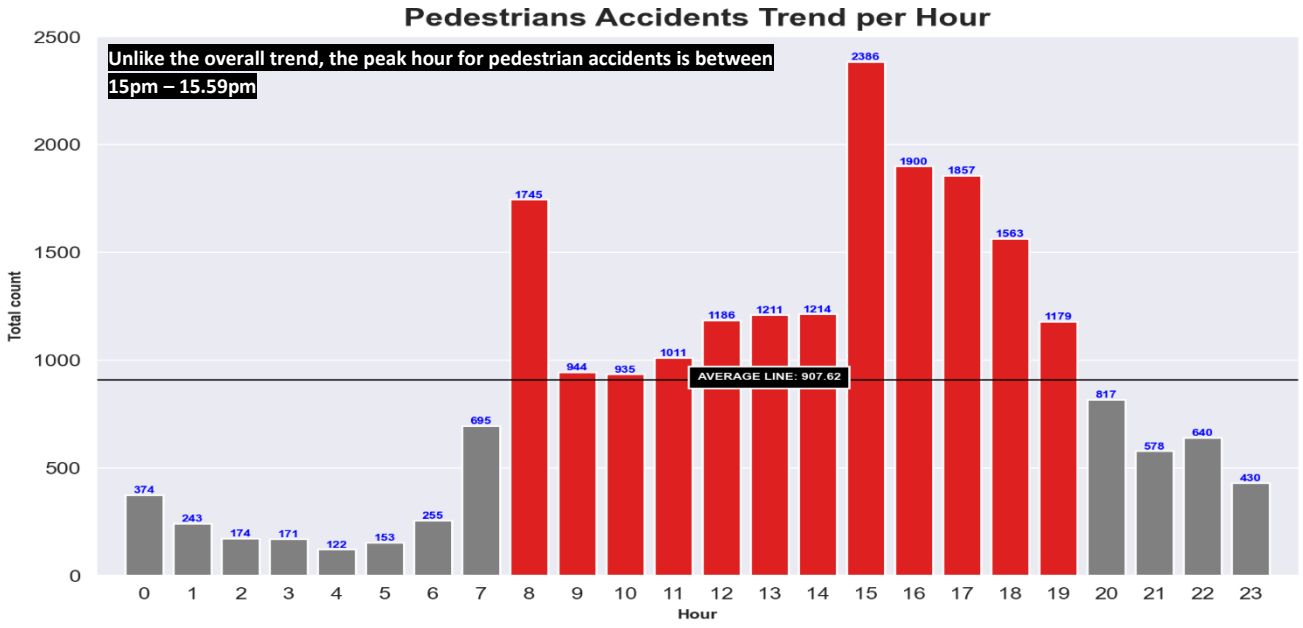


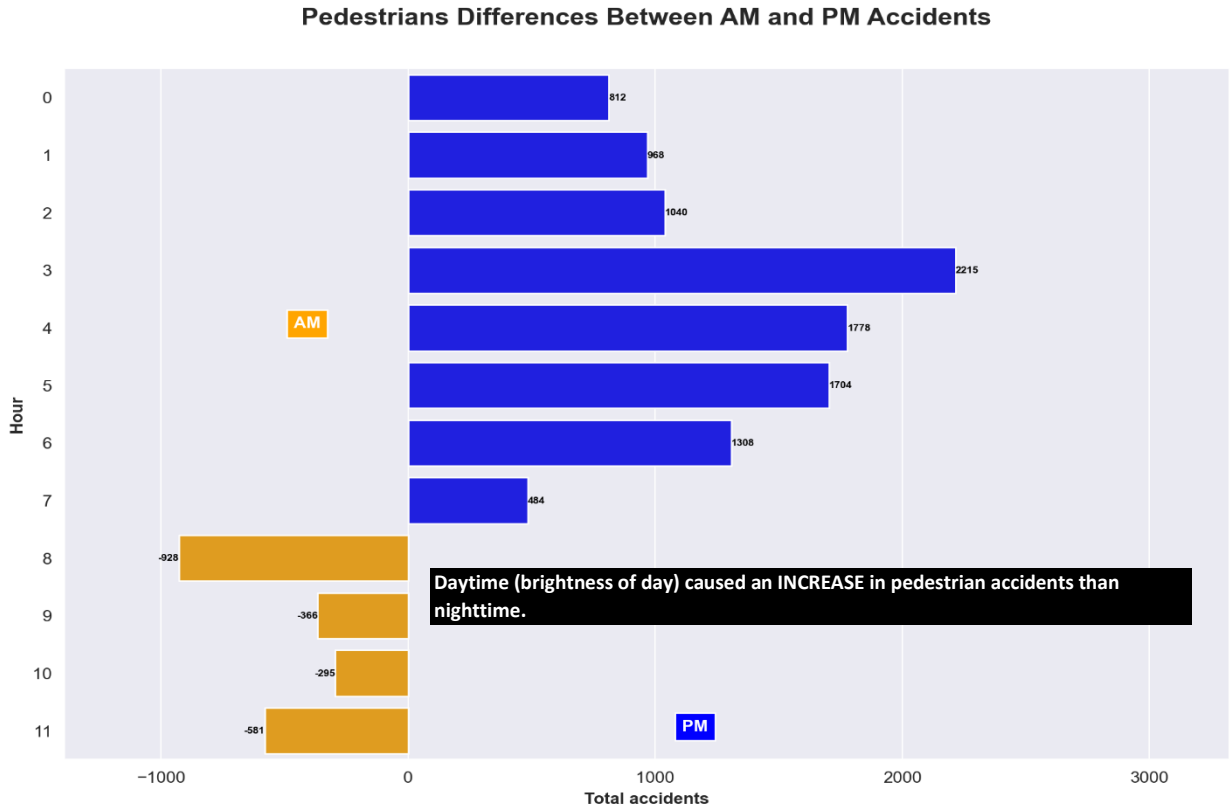





Monthly Pedestrian Accidents per Part of Day
January
February
March
April
May
June
July
August
September
October
November
December
279
364
348
374
289
205
139
257
321
224
239
212
192
201
229
263
238
235
293
280
239
217
266
241
198
302
192
137
210
241
231
304
355
223
190
208
214
220
242
292
245
285
280
320
311
266
228
200
158
188
171
191
254
266
300
238
334
233
279
278
298
234
266
233
290
277
325
261
301
190
287
308
317
328
424
327
207
334
298
266
293
307
258
225
month_name
0
50
100
150
200
250
300
350
400
total_count
day_name
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday

a. SIGNIFICANT HOURS OF THE DAY



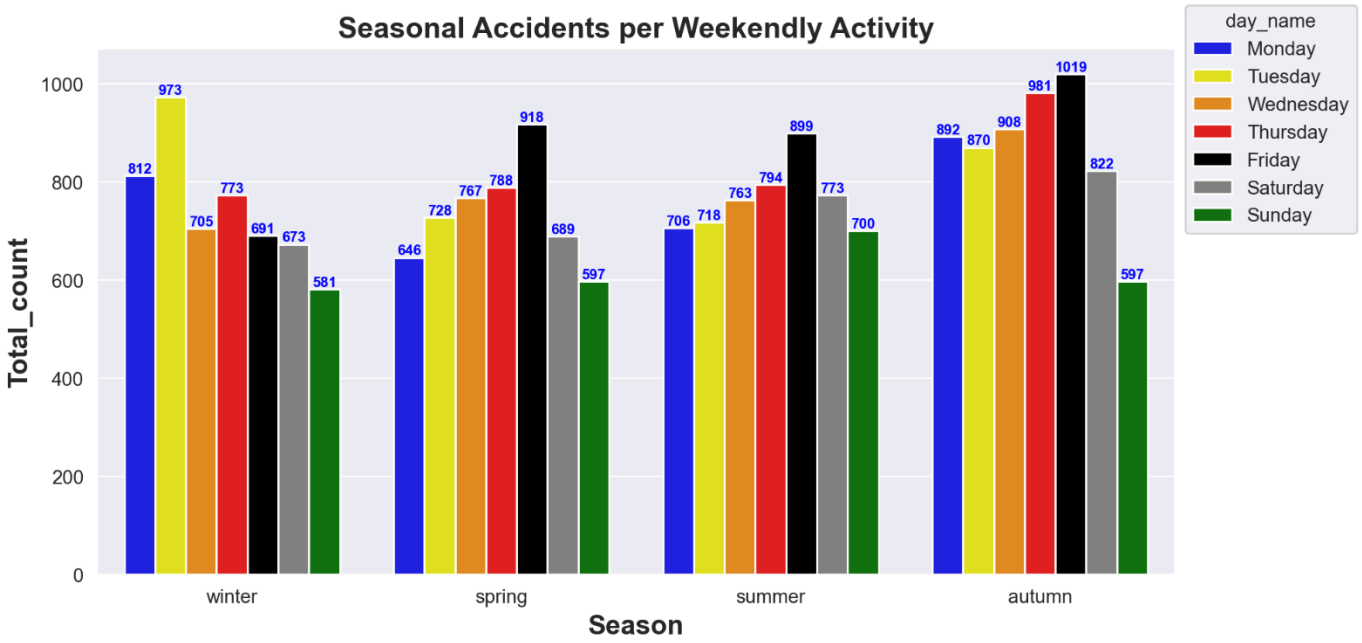


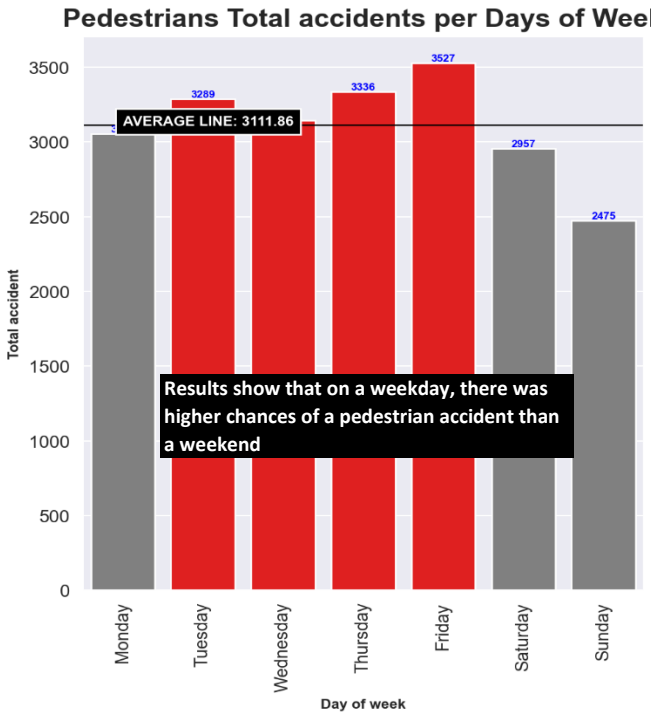
H0: Pedestrian accidents between 12pm and 18.59pm is not higher than between 0am and 6.59pm.

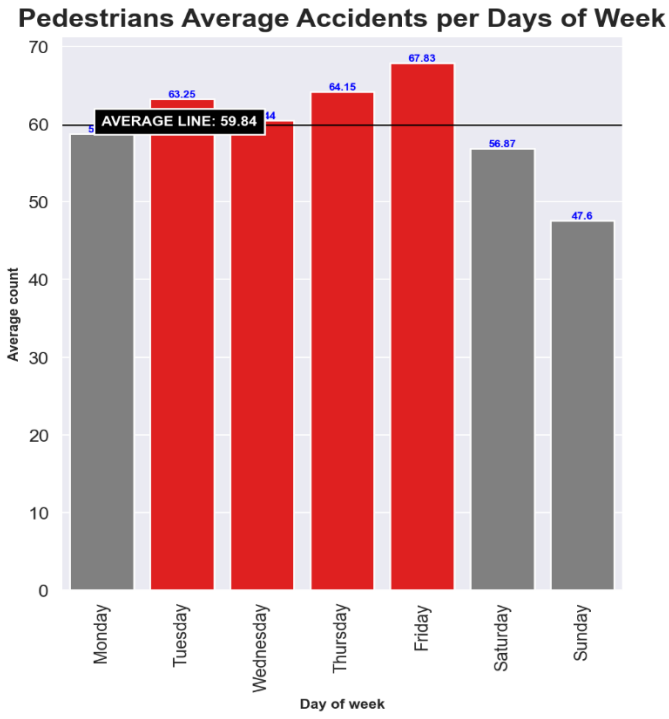
At 95% confidence, there is statistically significant evidence to show that H0 is false. Hence, we reject H0 and accept that hourly pedestrian accidents between 12pm and 18.59pm is significantly higher than between 0am and 6.59am.

Interpretation: These results suggest that 12pm - 18.59pm, and 8am – 11.59am were hours of significance, as it was more likely for a pedestrian accident in these periods than 0am – 6.59pm.

b. SIGNIFICANT DAYS OF THE WEEK



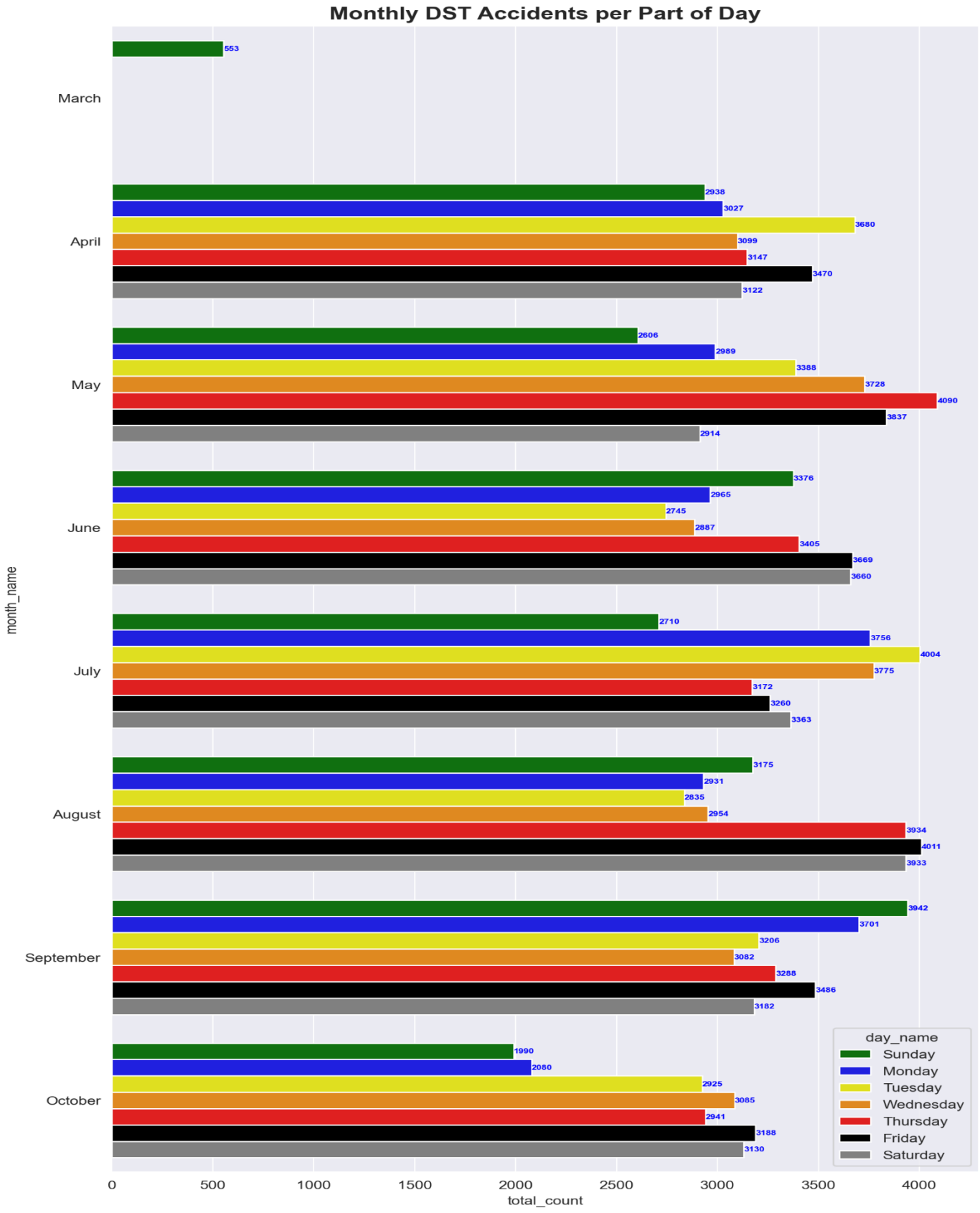




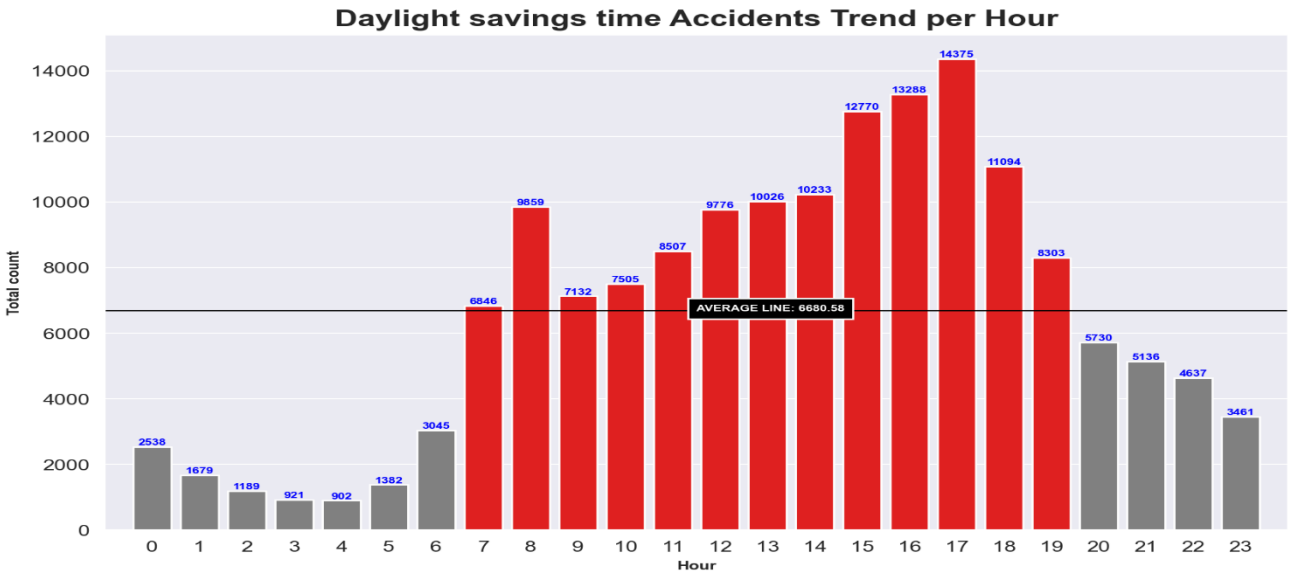
H0: Pedestrian accidents on weekdays were not higher than on weekends

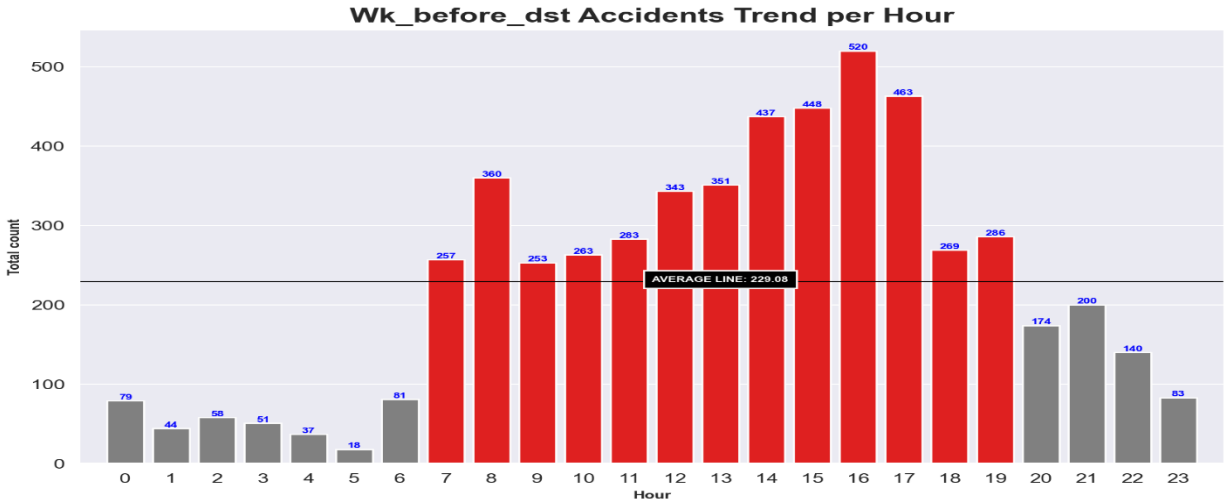
At 95% confidence, there is statistically significant evidence to show that H0 is false. Hence, we reject H0 and accept that pedestrian accidents were higher on weekdays than weekends.

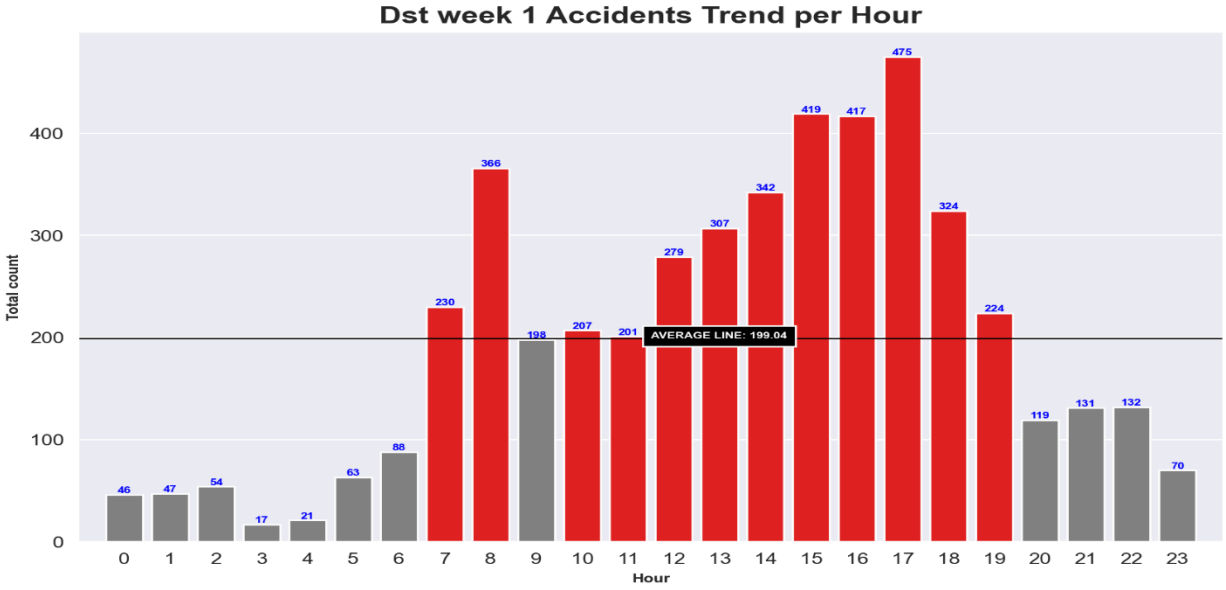
DAYLIGHT SAVING TIME (DST) IMPACT ON ACCIDENTS

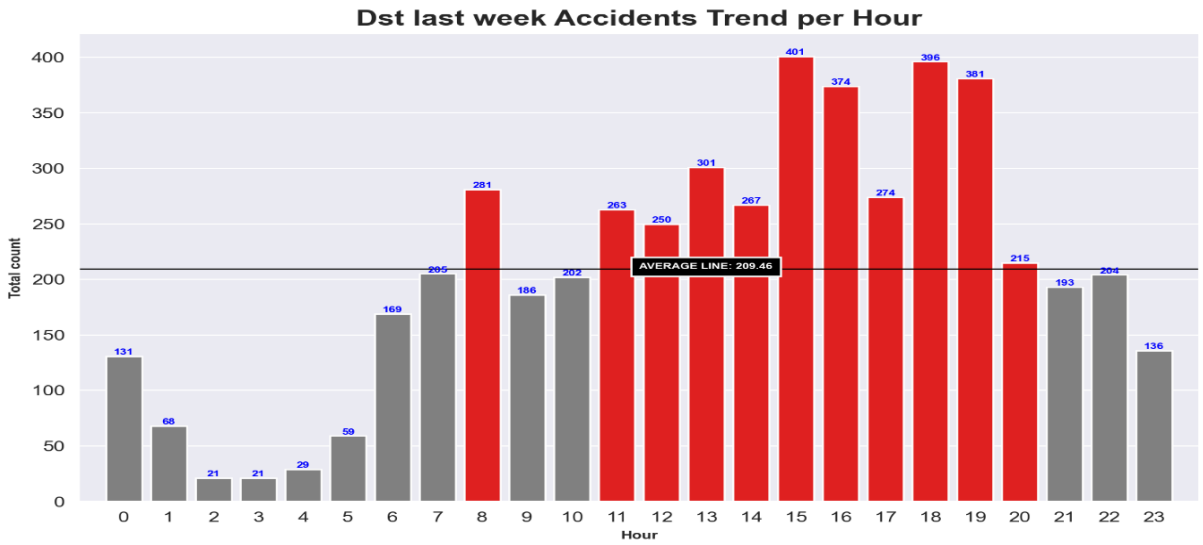


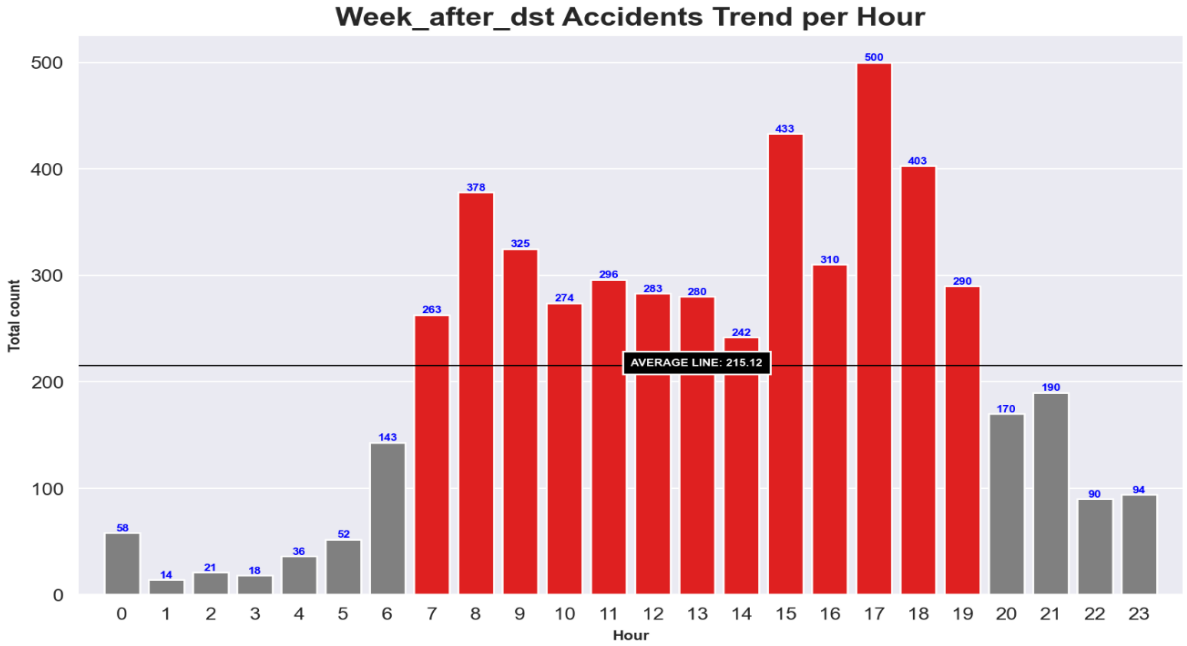
a. SIGNIFICANT HOURS OF THE DAY

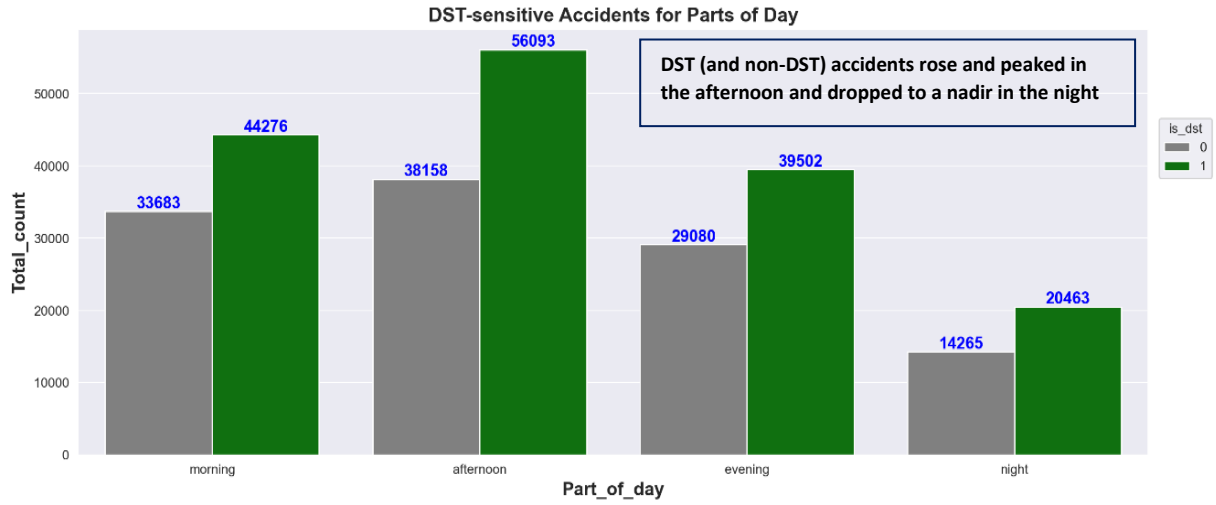




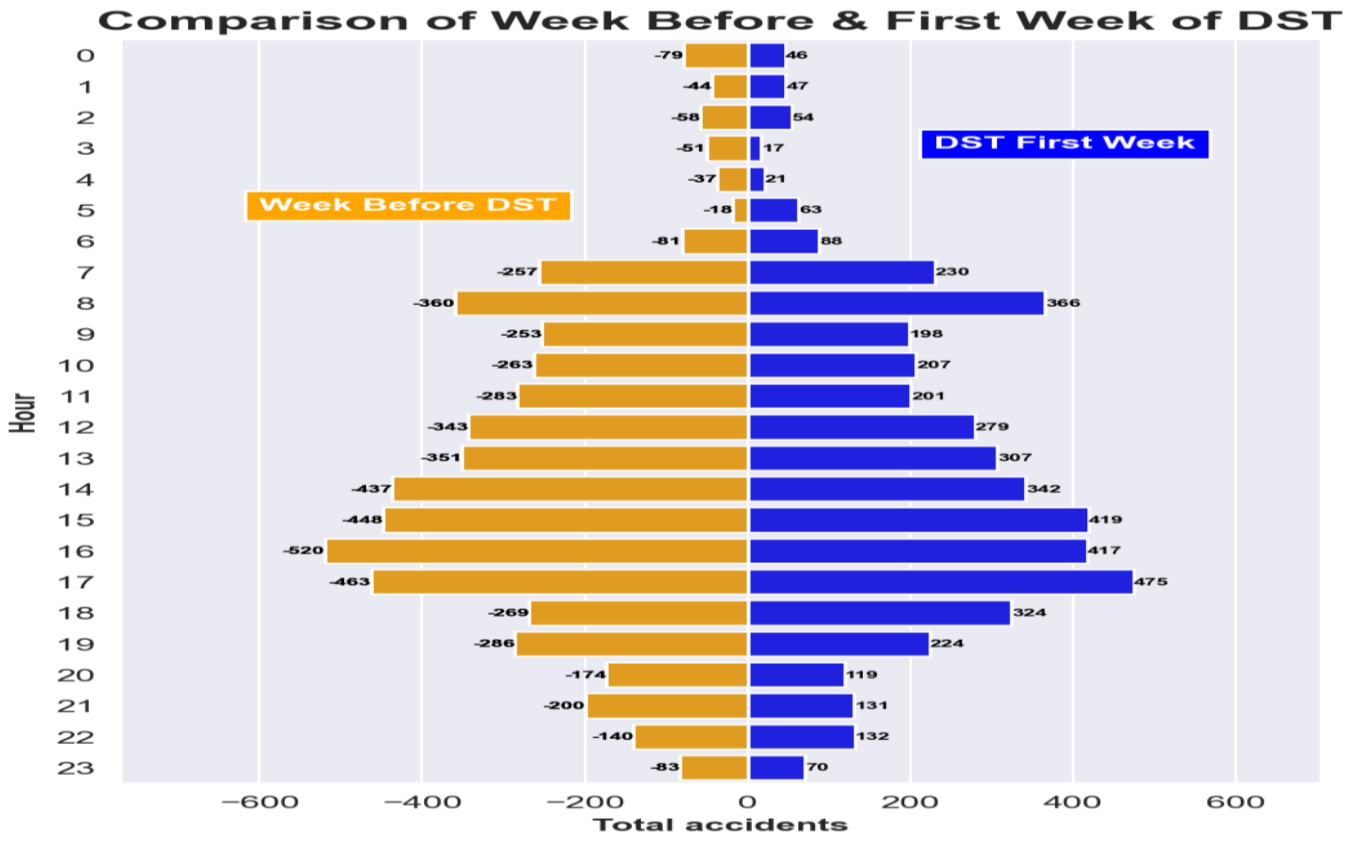


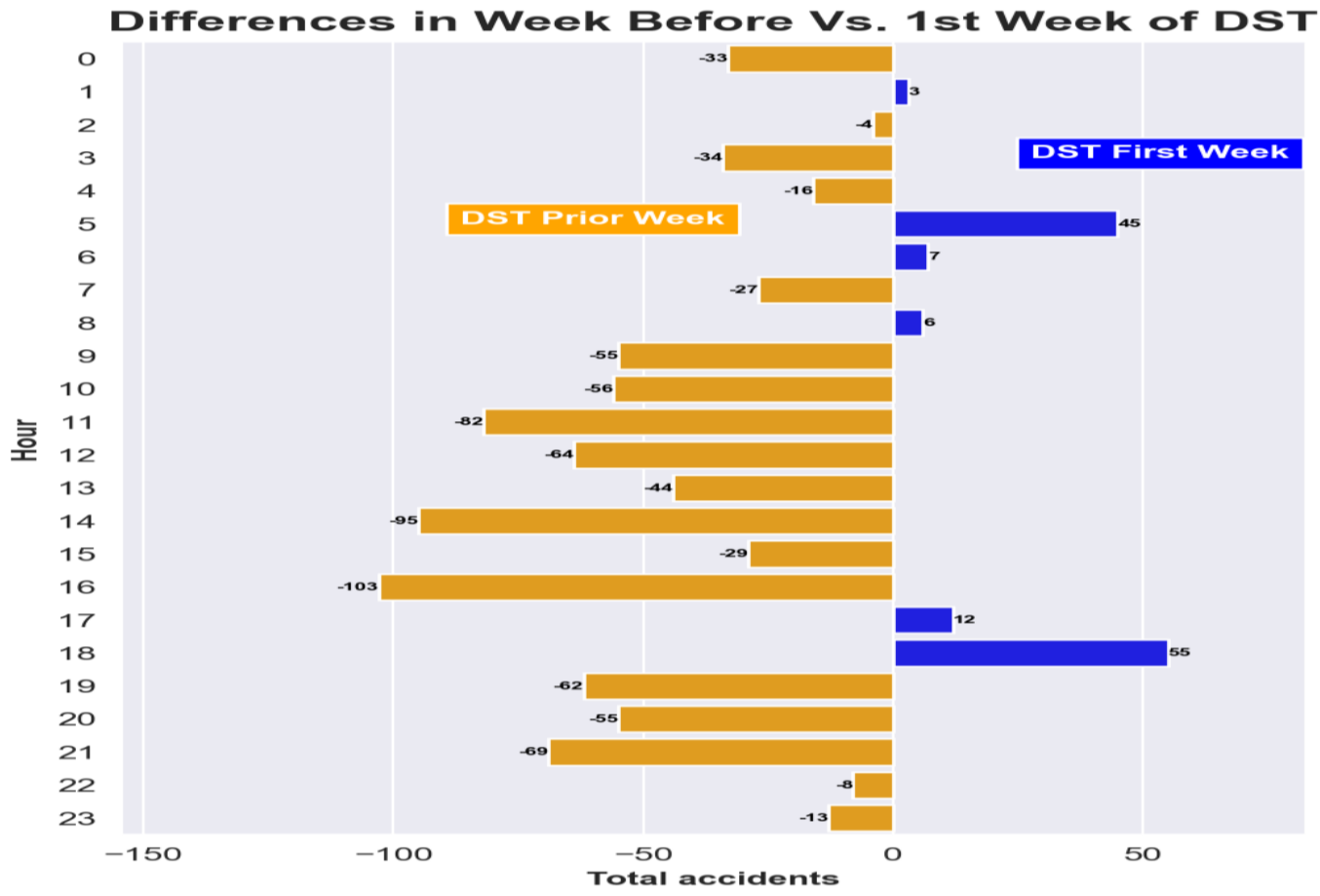

Dst last week Accidents Trend per Hour
Total count
Hour
AVERAGE LINE: 209.46
131
68
21
21
29
59
169
205
281
186
202
263
250
301
267
401
374
274
396
381
215
193
204
136


Week_after_dst Accidents Trend per Hour
Total count
Hour
AVERAGE LINE: 215.12
58
14
21
18
36
52
143
263
378
325
274
296
283
280
242
433
310
500
403
290
170
190
90
94


DST-sensitive Accidents for Parts of Day
DST (and non-DST) accidents rose and peaked in the afternoon and dropped to a nadir in the night
Total_count
Part_of_day
is_dst
0
1
33683
44276
38158
56093
29080
39502
14265
20463
morning
afternoon
evening
night

a. **COMPARING THE FIRST WEEK OF DST TO THE WEEK BEFORE DST**



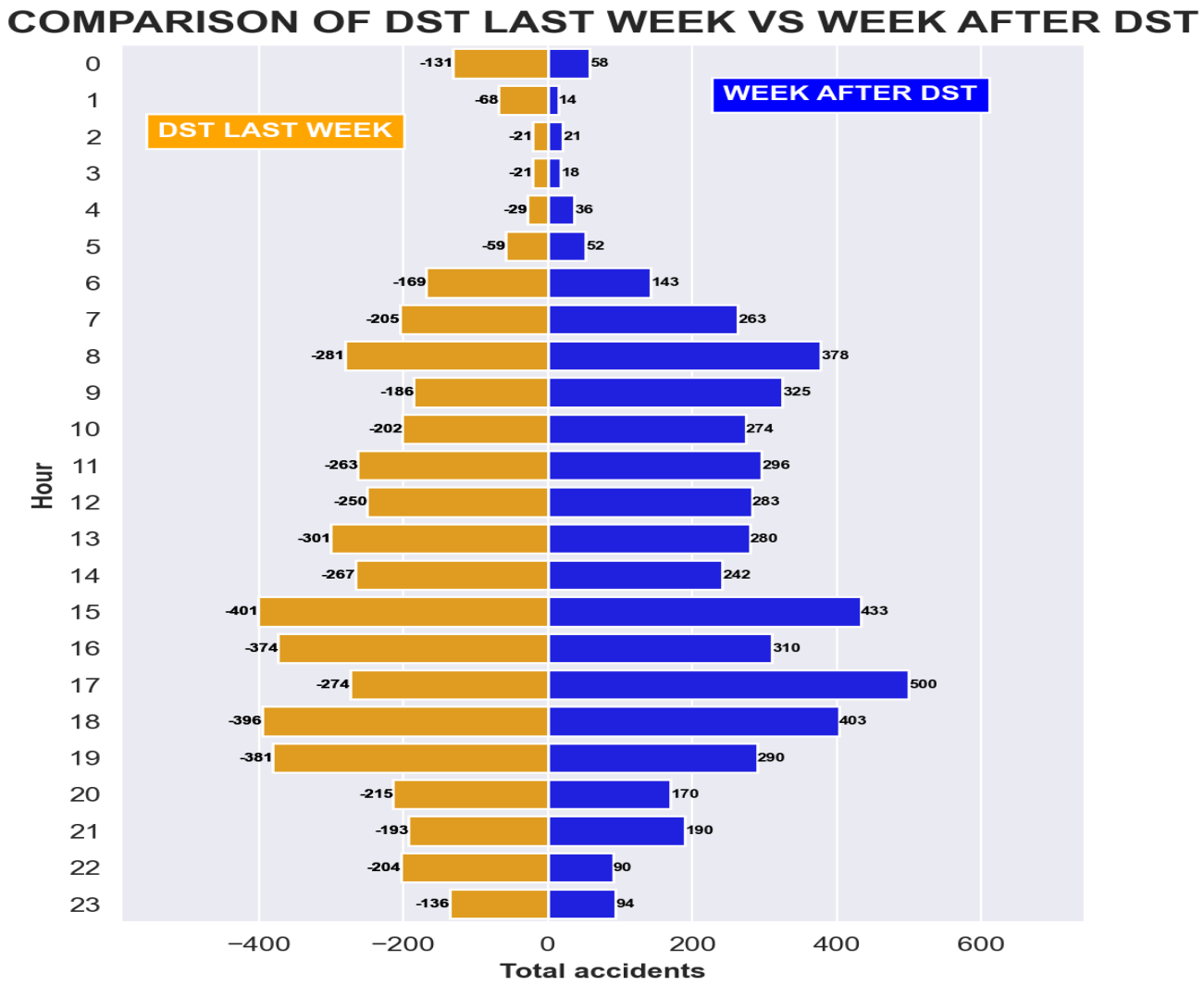


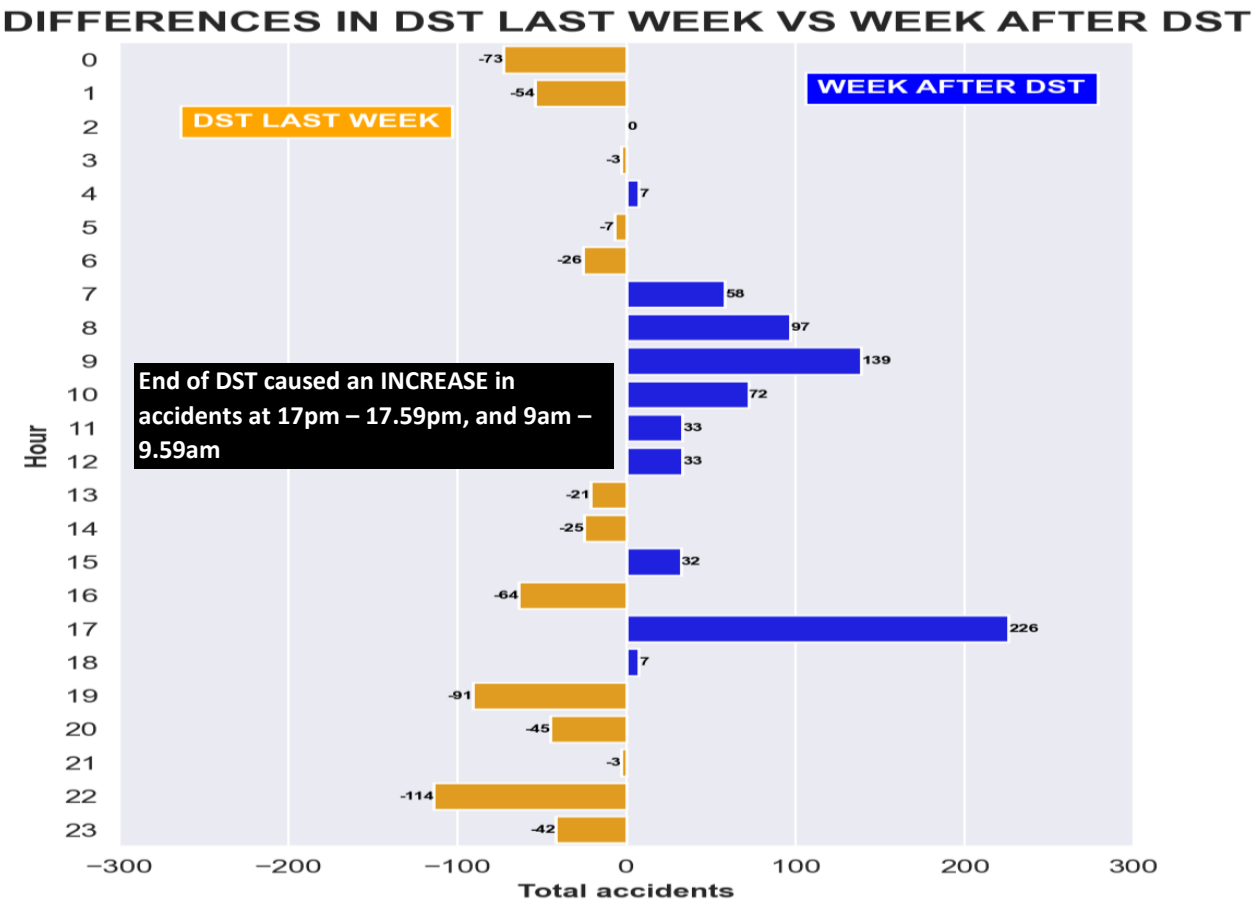
H0: 9am – 11.59am accidents in the week before DST were not higher than DST week 1

At 99% confidence, there is statistically significant evidence to show that H0 is false. Hence, we reject H0 and accept that 9am – 11.59am accidents in the week before were higher than DST week 1.

Interpretation: These results suggest that onset of DST caused a decrease in 9am – 11.59am accidents.

b. COMPARING LAST WEEK OF DST TO THE WEEK AFTER DST





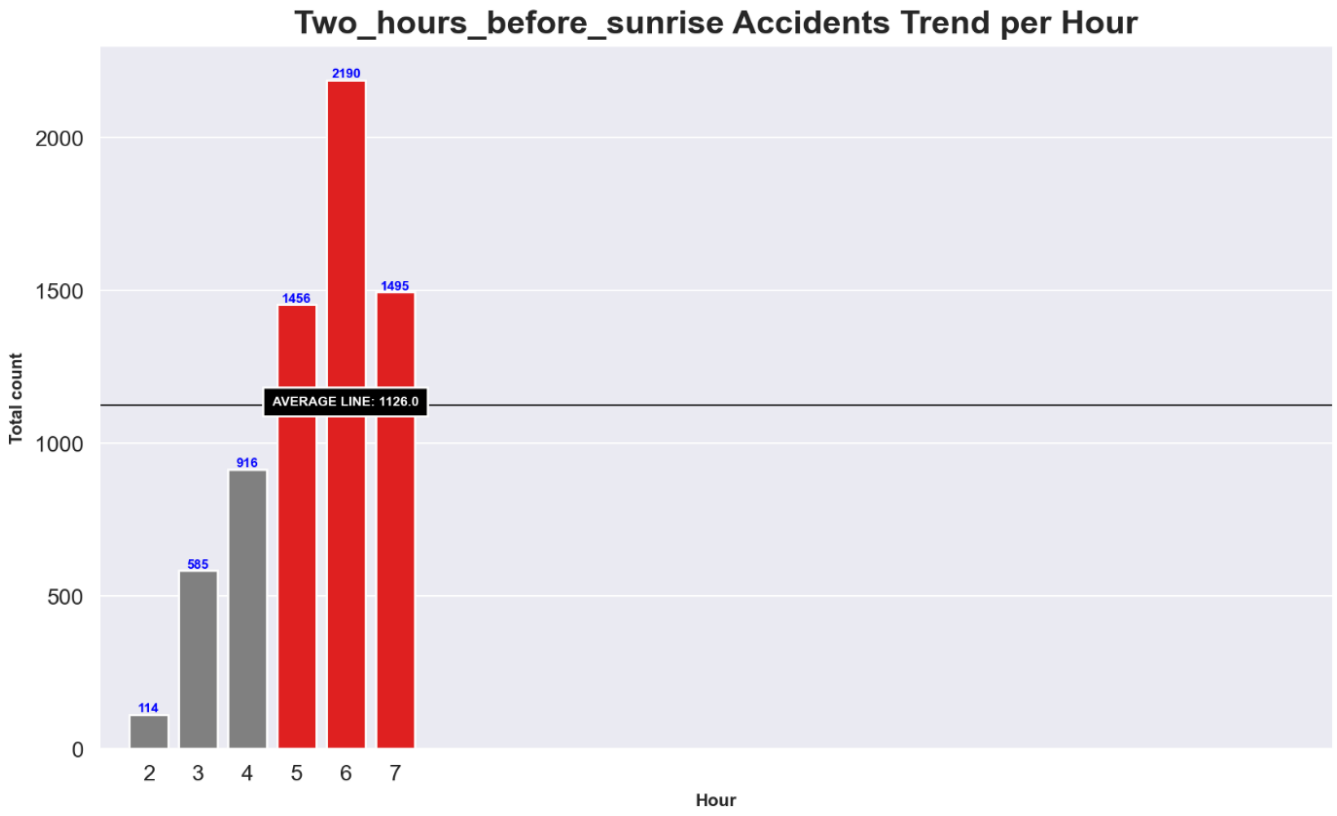
H0: 17pm - 17.59pm accidents in the week after DST were not higher than DST last week

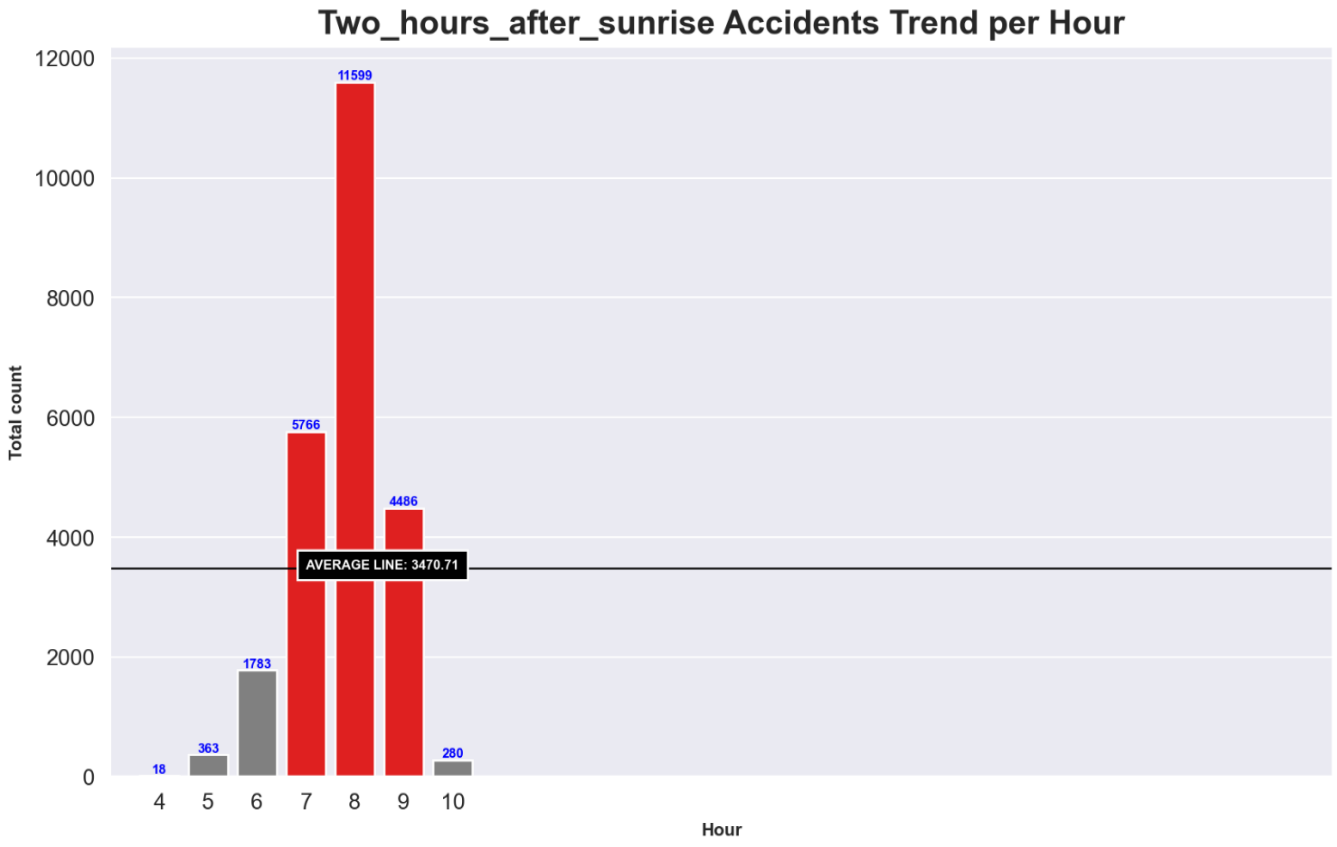
At 99% confidence, there is statistically significant evidence to show that H0 is false. Hence, we reject H0 and accept that 17pm - 17.59pm accidents in the week after DST were higher than the last week of DST.

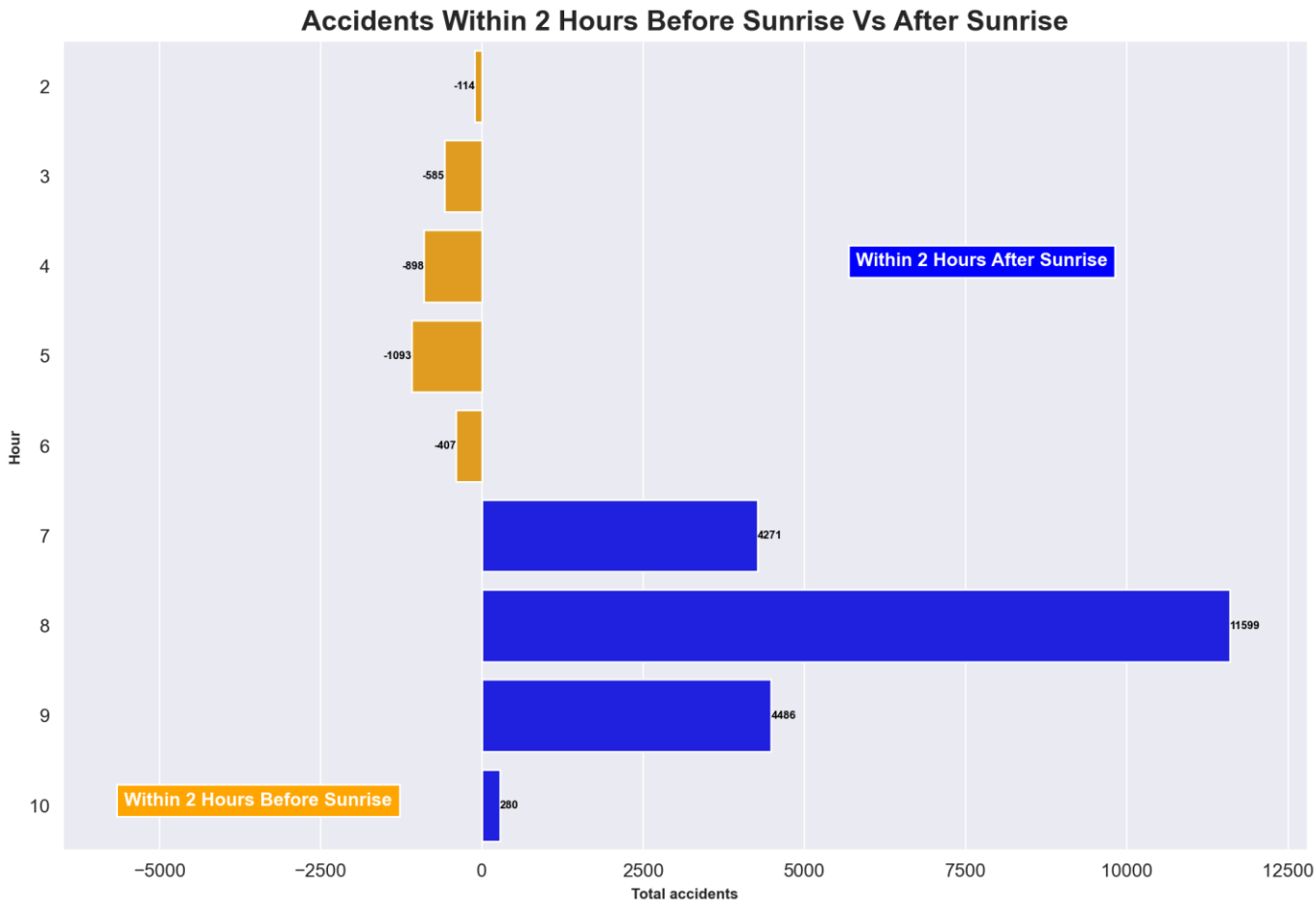
H0: 9am - 10.59am accidents in the week after DST were not higher the last week of DST

At 95% confidence, there is statistically significant evidence to show that H0 is false. Hence, we reject H0 and accept that 9am - 10.59pm accidents in the week after DST were higher than the last week of DST.

IMPACT OF SUNRISE & SUNSET TIMES



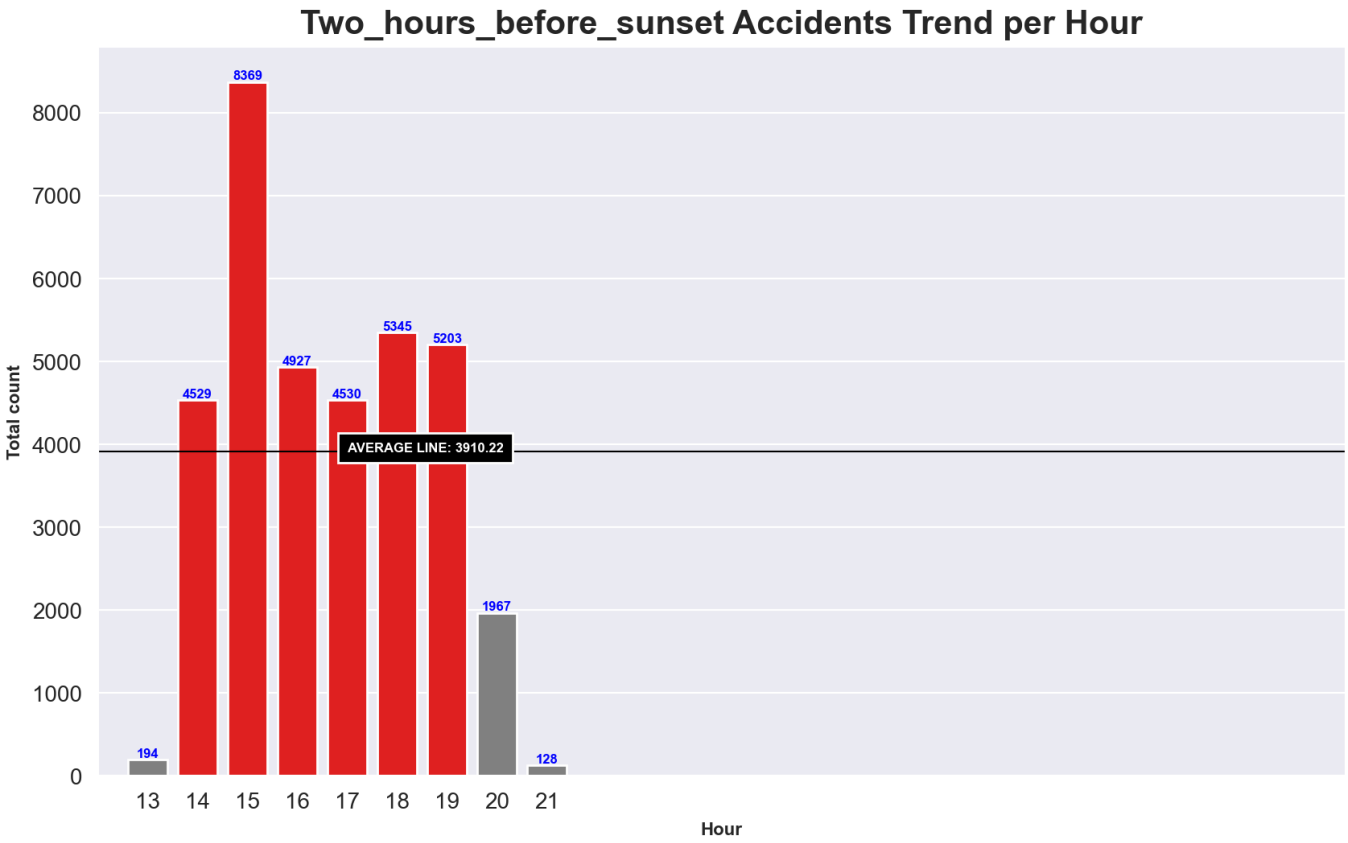


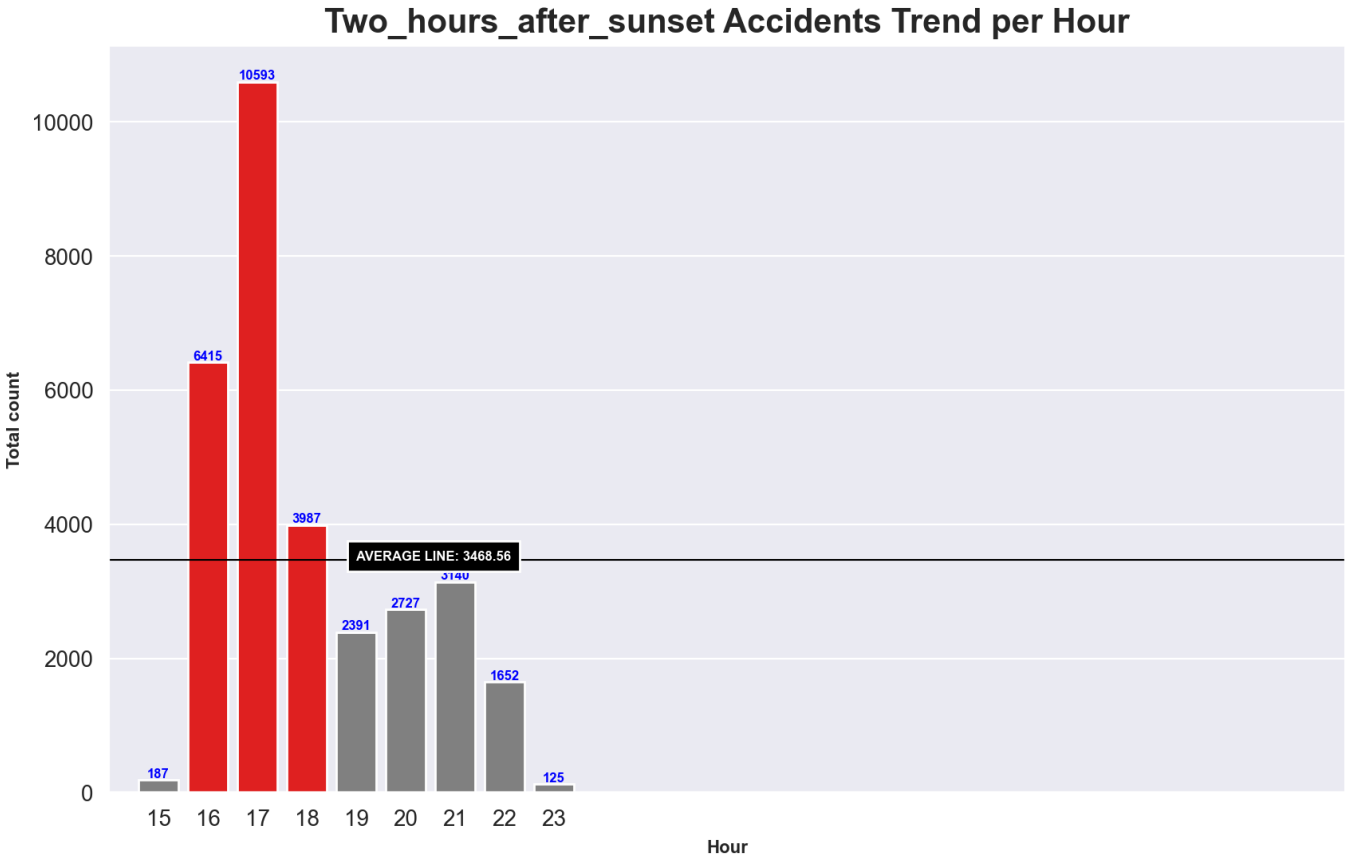


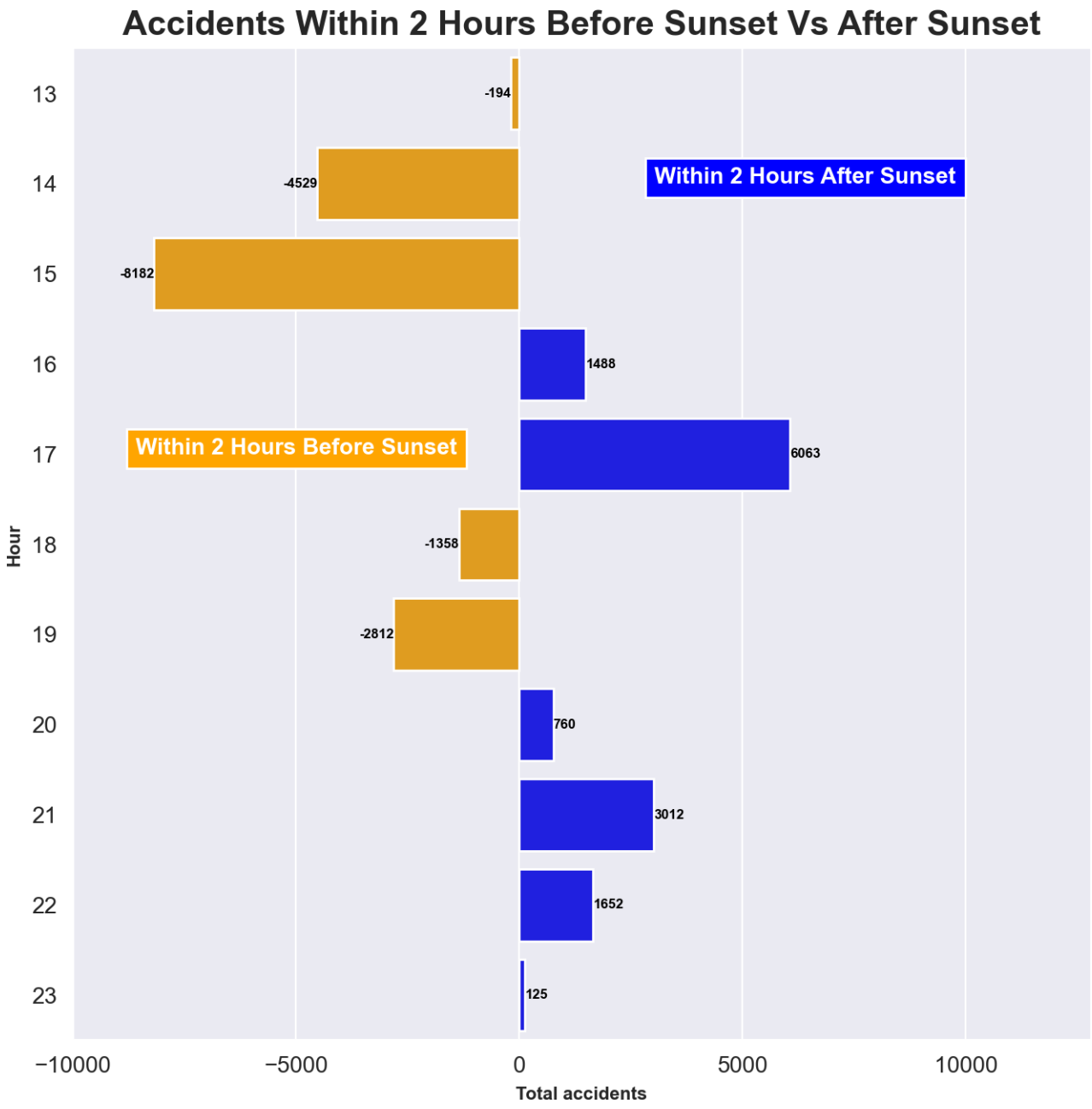
H0: Accidents within 2 hours after sunrise at 7am - 9.59am were not higher than within two hours before sunrise

At 95% confidence, there is evidence to show that H0 is false. Hence, we reject H0 and accept that accidents within two hours after sunrise at 7am – 9.59am were higher than within two hours before sunrise.

Interpretation: These results suggest that sunrise caused an increase in accidents at 7am – 9.59am.


Two_hours_before_sunset Accidents Trend per Hour
Total count
Hour
AVERAGE LINE: 3910.22
194
4529
8369
4927
4530
5345
5203
1967
128
0
1000
2000
3000
4000
5000
6000
7000
8000
13
14
15
16
17
18
19
20
21


Two_hours_after_sunset Accidents Trend per Hour
Total count
Hour
AVERAGE LINE: 3468.56
187
6415
10593
3987
2391
2727
3140
1652
125
0
2000
4000
6000
8000
10000
15
16
17
18
19
20
21
22
23

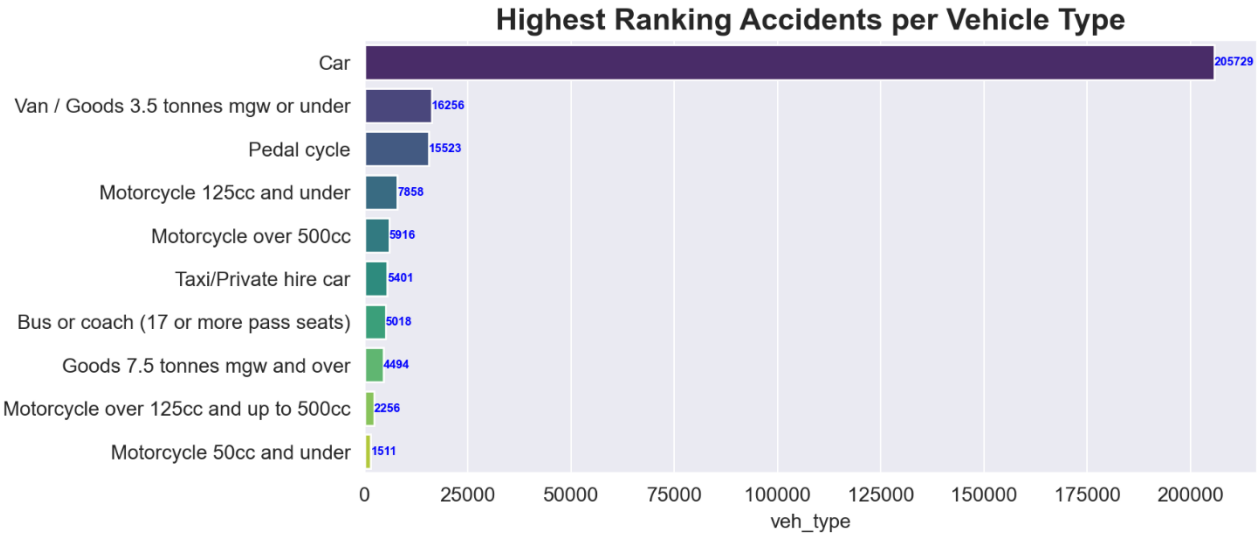


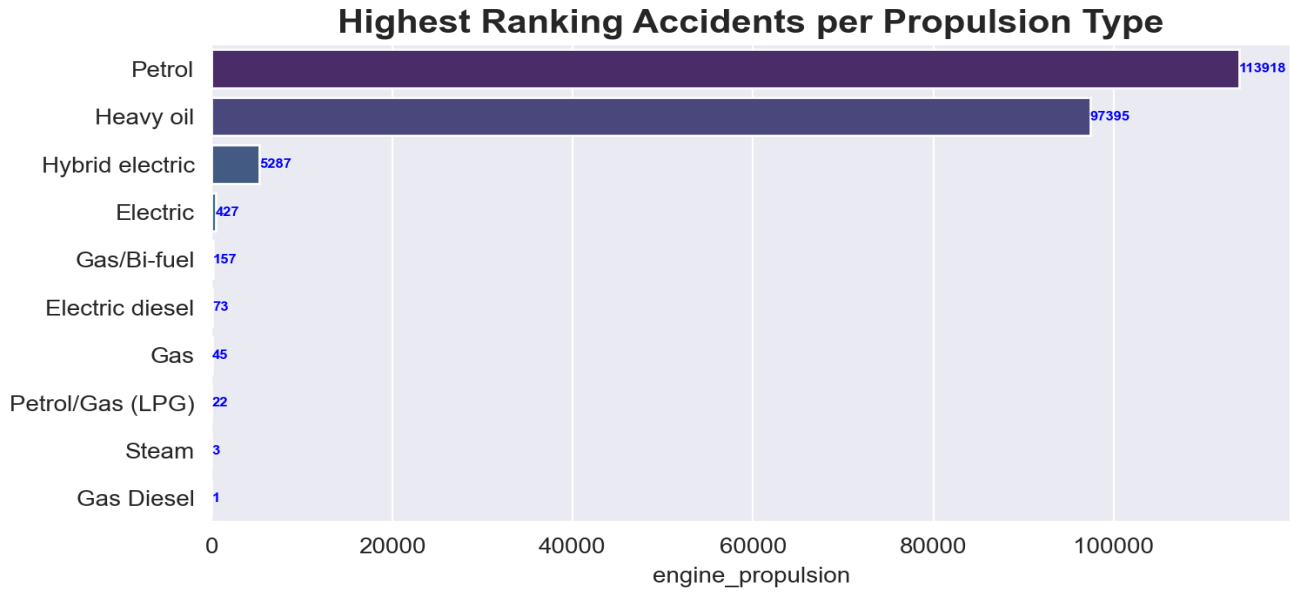
H0: Accidents within two hours before sunset at 14 - 15.59pm is not higher than after sunset

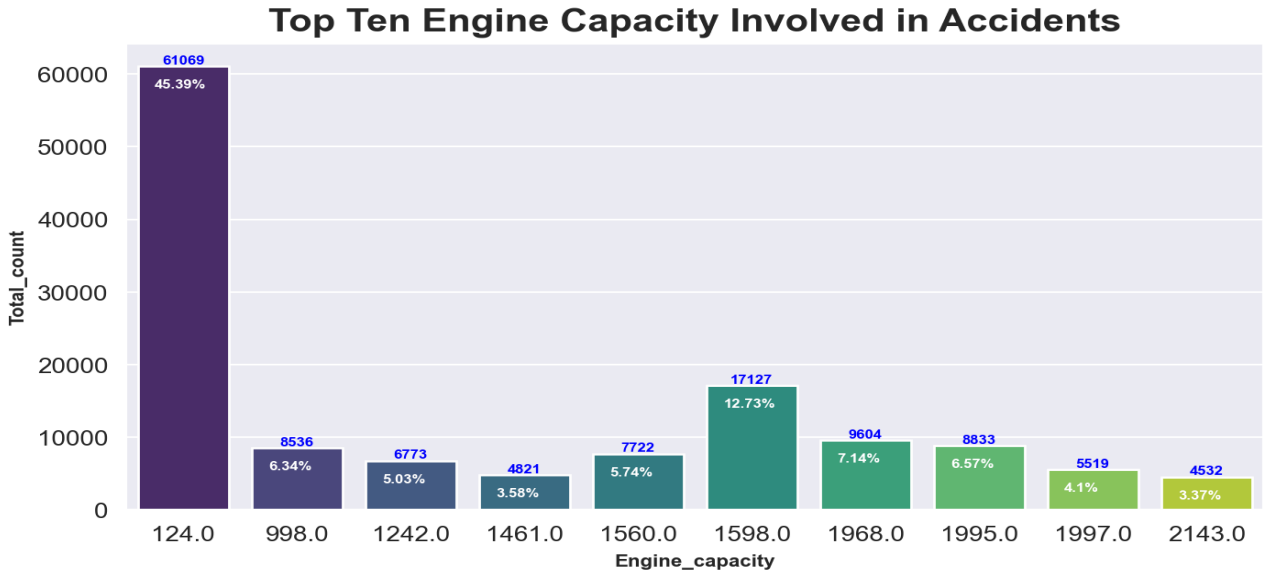
At 99% confidence, there is evidence to show that H0 is false. We reject H0 and accept that accidents within two hours before sunset at 14pm – 15.59pm were higher than within two hours before sunset.

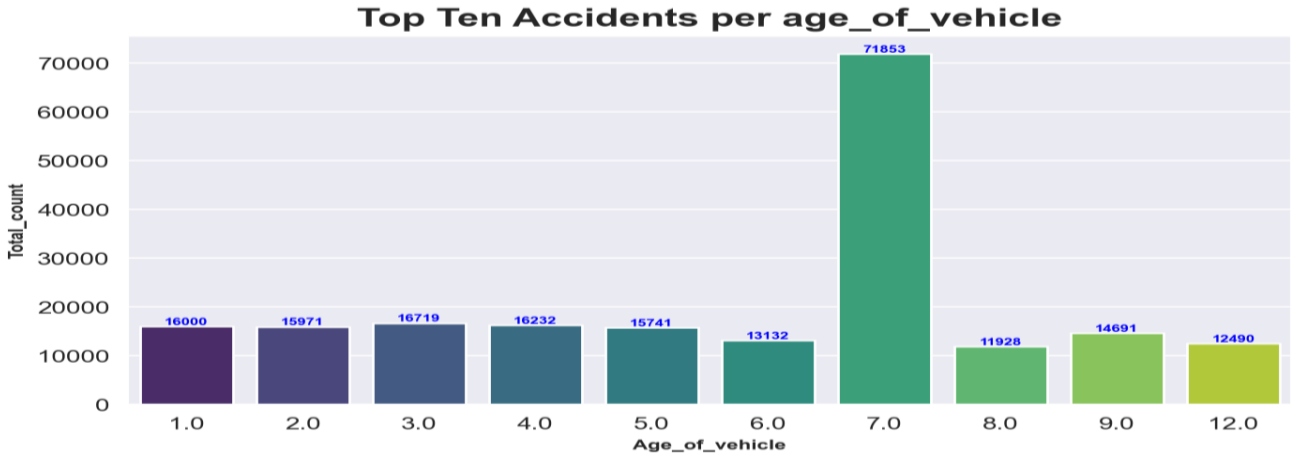
Interpretation: These results suggest that sunset caused a decline in accidents at 14pm – 15.59pm.

VEHICULAR VARIABLE INFLUENCE

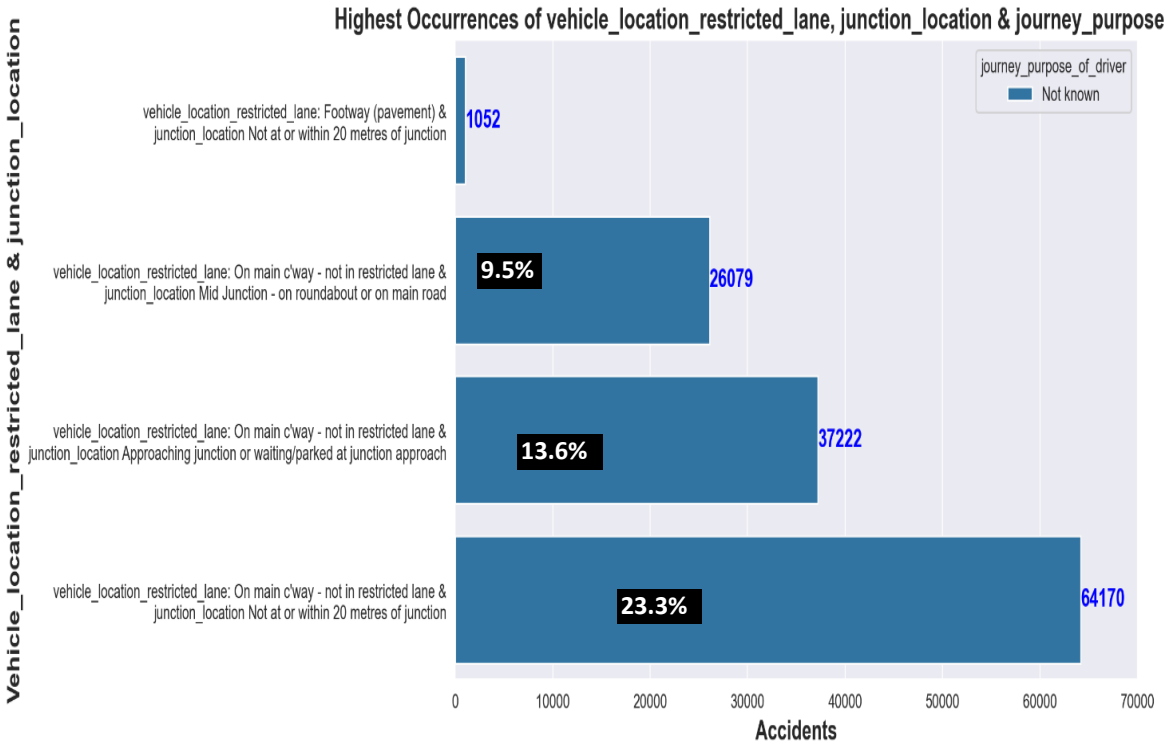


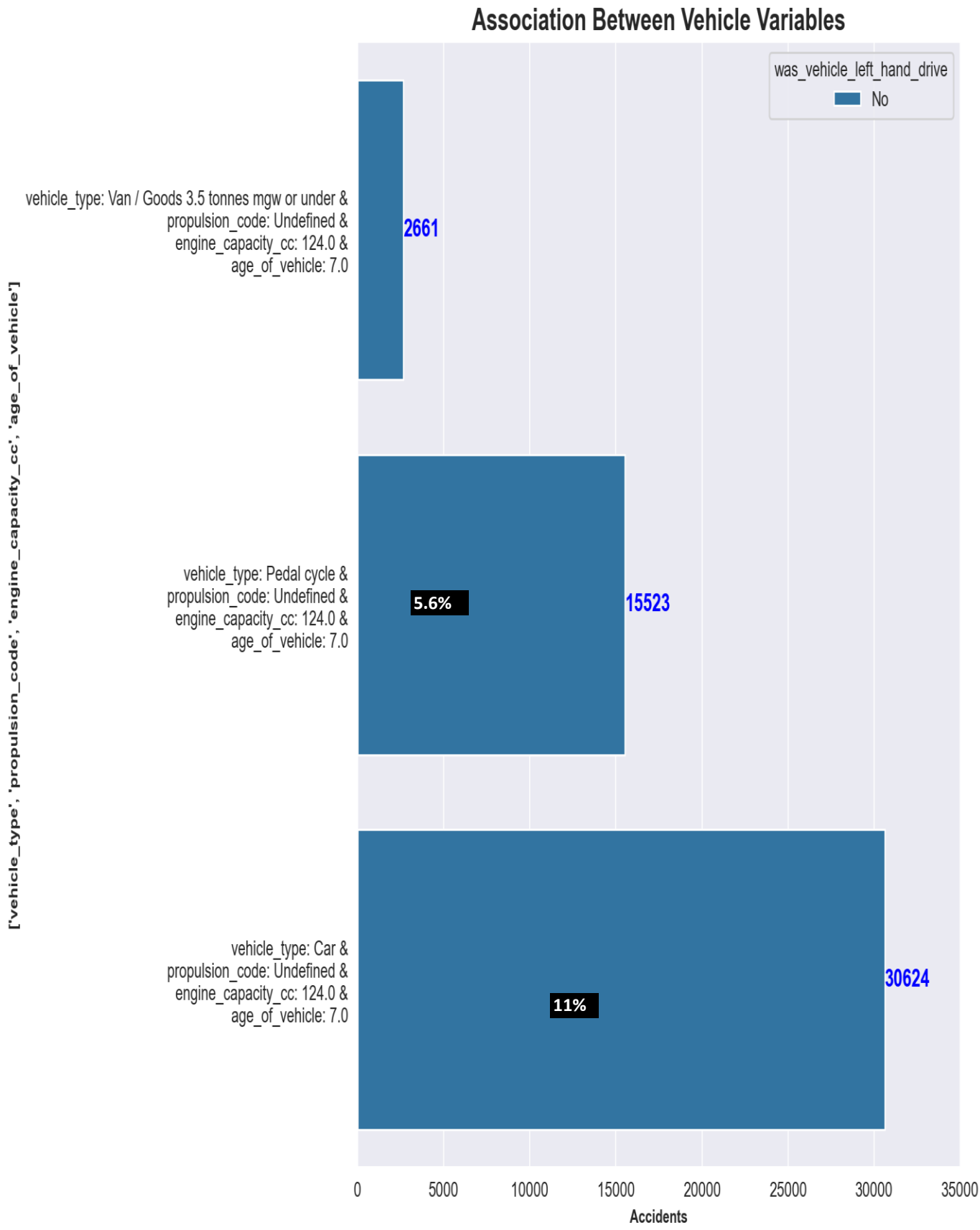




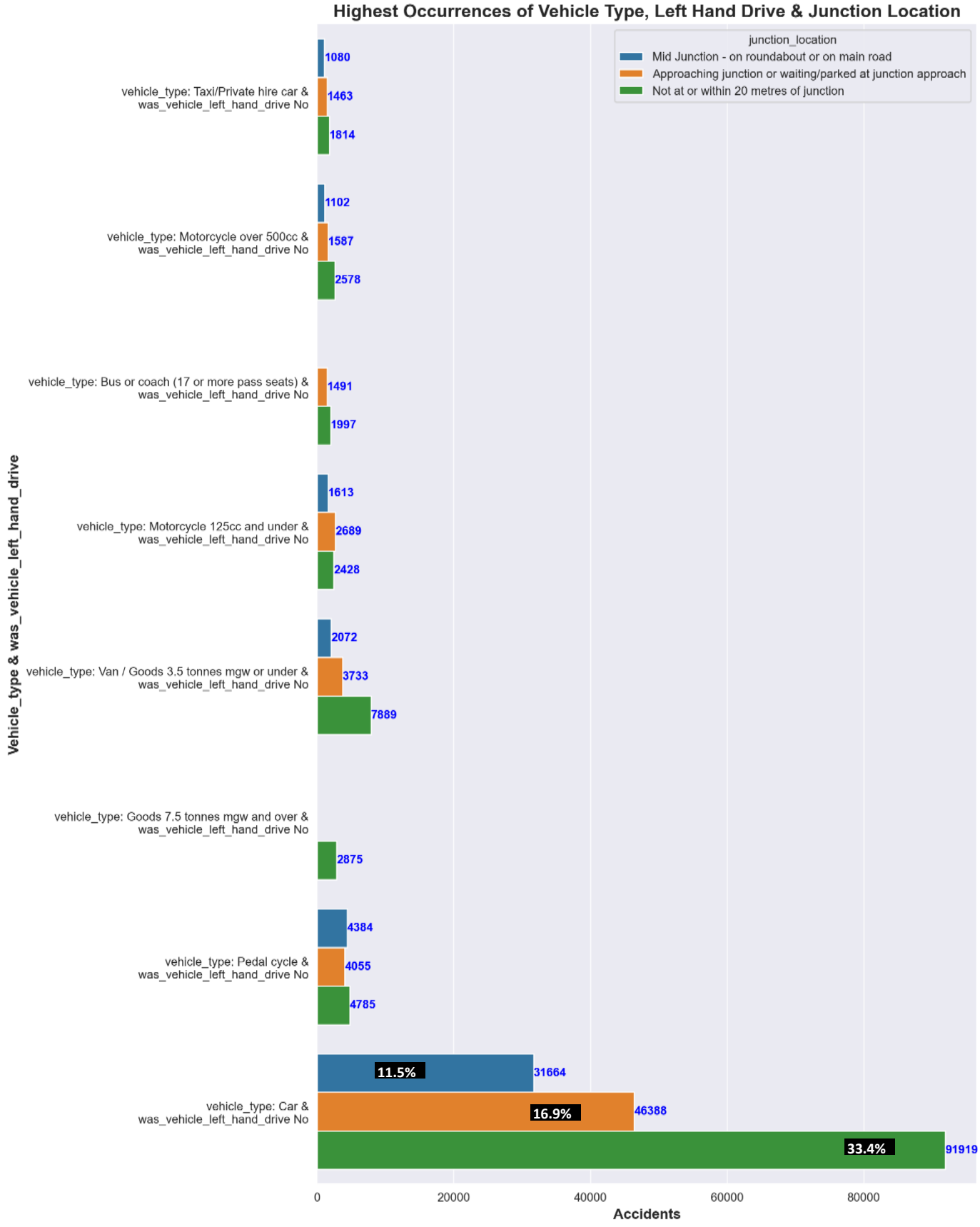

Top Ten Accidents per age_of_vehicle
Total_count
70000
60000
50000
40000
30000
20000
10000
0
16000
15971
16719
16232
15741
13132
71853
11928
14691
12490
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
12.0
Age_of_vehicle

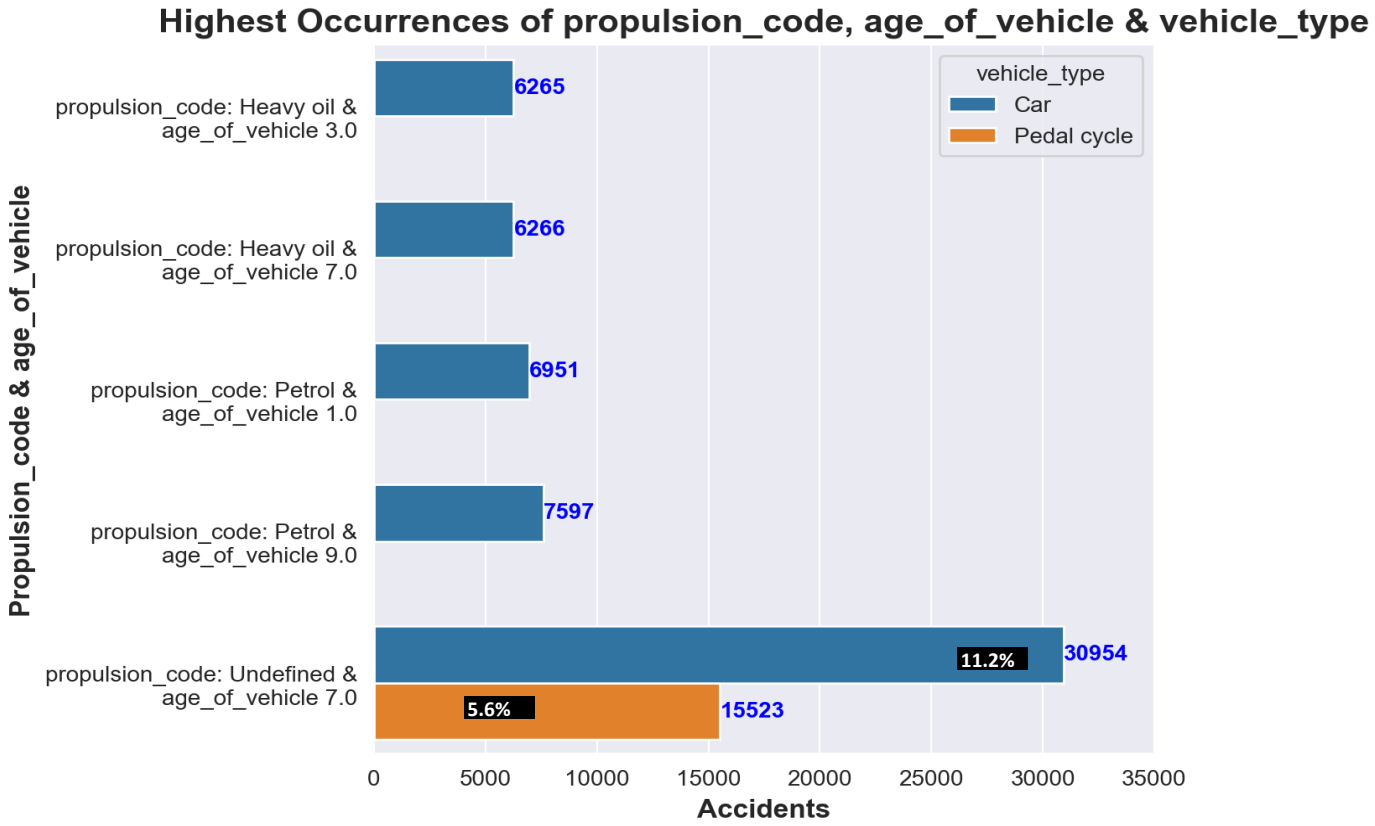

Highest Ranking Accidents per Lefthandedness
Total_count
250000
200000
150000
100000
50000
0
273846
99.39%
1674
0.61%
No
Yes
Left_hand_drive

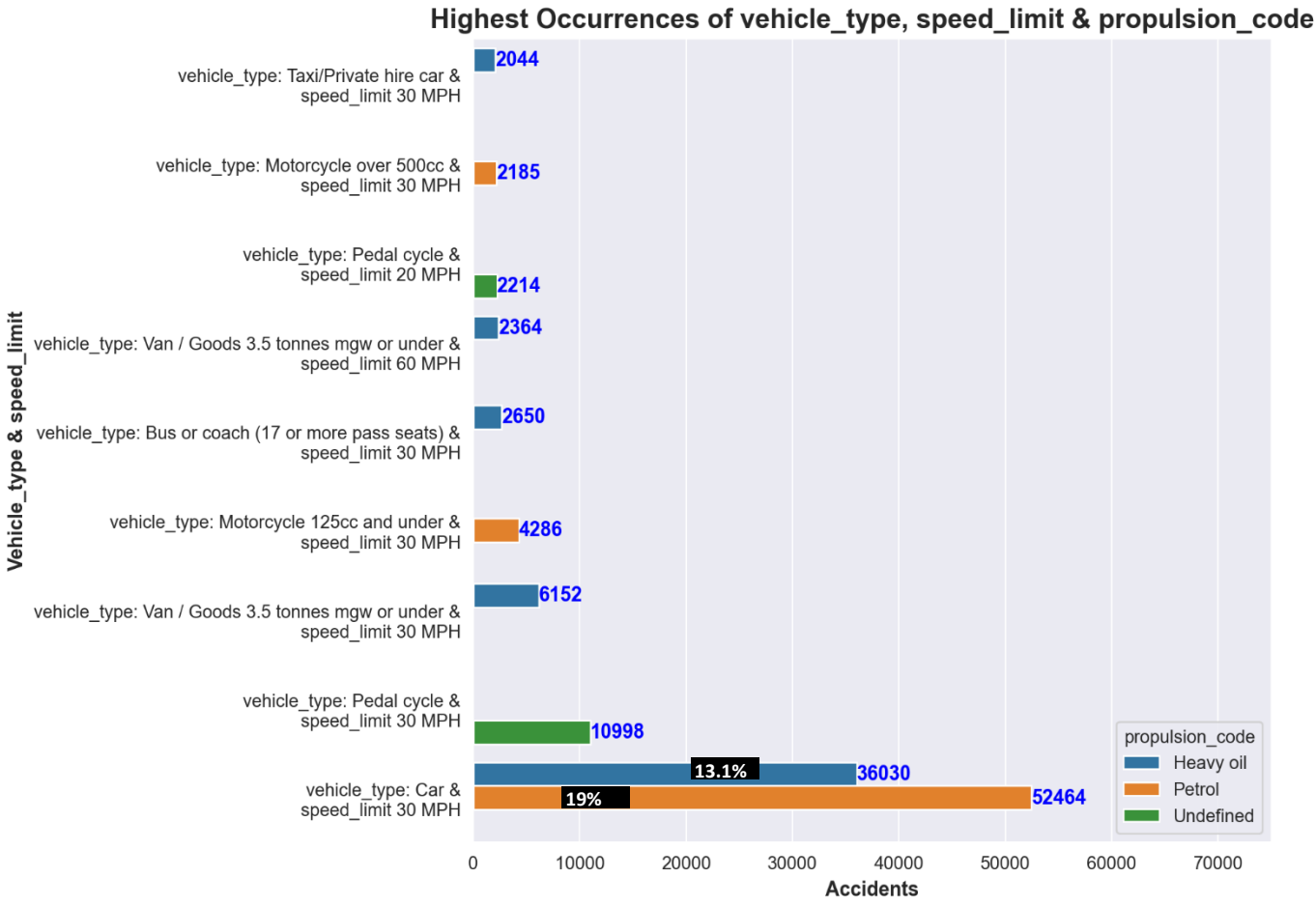

Highest Occurrences of vehicle_location_restricted_lane, junction_location & journey_purpose_of_driver
Vehicle_location_restricted_lane & junction_location
journey_purpose_of_driver
Not known
vehicle_location_restricted_lane: Footway (pavement) & junction_location Not at or within 20 metres of junction
1052
vehicle_location_restricted_lane: On main c'way - not in restricted lane & junction_location Mid Junction - on roundabout or on main road
9.5%
26079
vehicle_location_restricted_lane: On main c'way - not in restricted lane & junction_location Approaching junction or waiting/parked at junction approach
13.6%
37222
vehicle_location_restricted_lane: On main c'way - not in restricted lane & junction_location Not at or within 20 metres of junction
23.3%
64170
0
10000
20000
30000
40000
50000
60000
70000
Accidents


Association Between Vehicle Variables
was_vehicle_left_hand_drive
No
['vehicle_type', 'propulsion_code', 'engine_capacity_cc', 'age_of_vehicle']
vehicle_type: Van / Goods 3.5 tonnes mgw or under & propulsion_code: Undefined & engine_capacity_cc: 124.0 & age_of_vehicle: 7.0
2661
vehicle_type: Pedal cycle & propulsion_code: Undefined & engine_capacity_cc: 124.0 & age_of_vehicle: 7.0
5.6%
15523
vehicle_type: Car & propulsion_code: Undefined & engine_capacity_cc: 124.0 & age_of_vehicle: 7.0
11%
30624
0
5000
10000
15000
20000
25000
30000
35000
Accidents

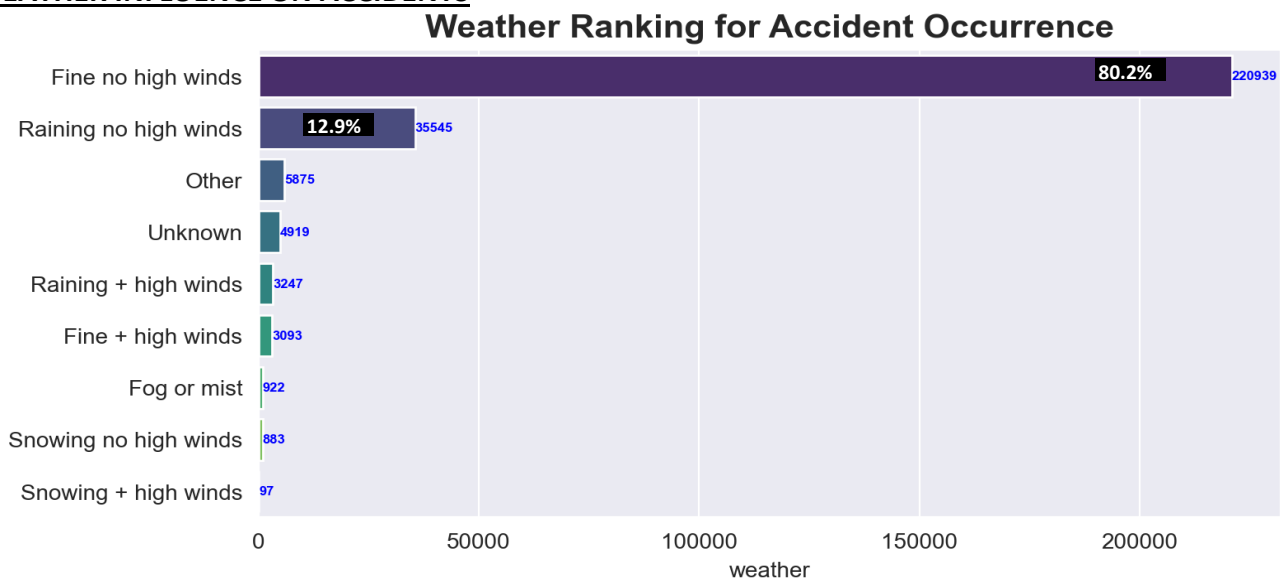

Highest Occurrences of vehicle_manoeuvre, vehicle_location_restricted_lane & propulsion_code
Vehicle_manoeuvre & vehicle_location_restricted_lane
vehicle_manoeuvre: Reversing & vehicle_location_restricted_lane On main c'way - not in restricted lane
1191
vehicle_manoeuvre: Overtaking static vehicle - offside & vehicle_location_restricted_lane On main c'way - not in restricted lane
1220
vehicle_manoeuvre: Overtaking moving vehicle - offside & vehicle_location_restricted_lane On main c'way - not in restricted lane
1377
2463
vehicle_manoeuvre: Going ahead left-hand bend & vehicle_location_restricted_lane On main c'way - not in restricted lane
2882
4084
1418
vehicle_manoeuvre: Waiting to turn right & vehicle_location_restricted_lane On main c'way - not in restricted lane
1489
1791
vehicle_manoeuvre: Going ahead right-hand bend & vehicle_location_restricted_lane On main c'way - not in restricted lane
3199
4361
1587
vehicle_manoeuvre: Turning left & vehicle_location_restricted_lane On main c'way - not in restricted lane
2644
3154
1672
vehicle_manoeuvre: Waiting to go - held up & vehicle_location_restricted_lane On main c'way - not in restricted lane
6232
6606
2132
vehicle_manoeuvre: Moving off & vehicle_location_restricted_lane On main c'way - not in restricted lane
3966
4695
2263
vehicle_manoeuvre: Parked & vehicle_location_restricted_lane On main c'way - not in restricted lane
3766
3782
2451
vehicle_manoeuvre: Slowing or stopping & vehicle_location_restricted_lane On main c'way - not in restricted lane
8291
8070
2923
vehicle_manoeuvre: Turning right & vehicle_location_restricted_lane On main c'way - not in restricted lane
8418
11061
4150
vehicle_manoeuvre: Going ahead other & vehicle_location_restricted_lane On main c'way - not in restricted lane
17.4%
47903
20.3%
56054
11.5%
31661
propulsion_code
Heavy oil
Petrol
Undefined
0
10000
20000
30000
40000
50000
60000
Accidents

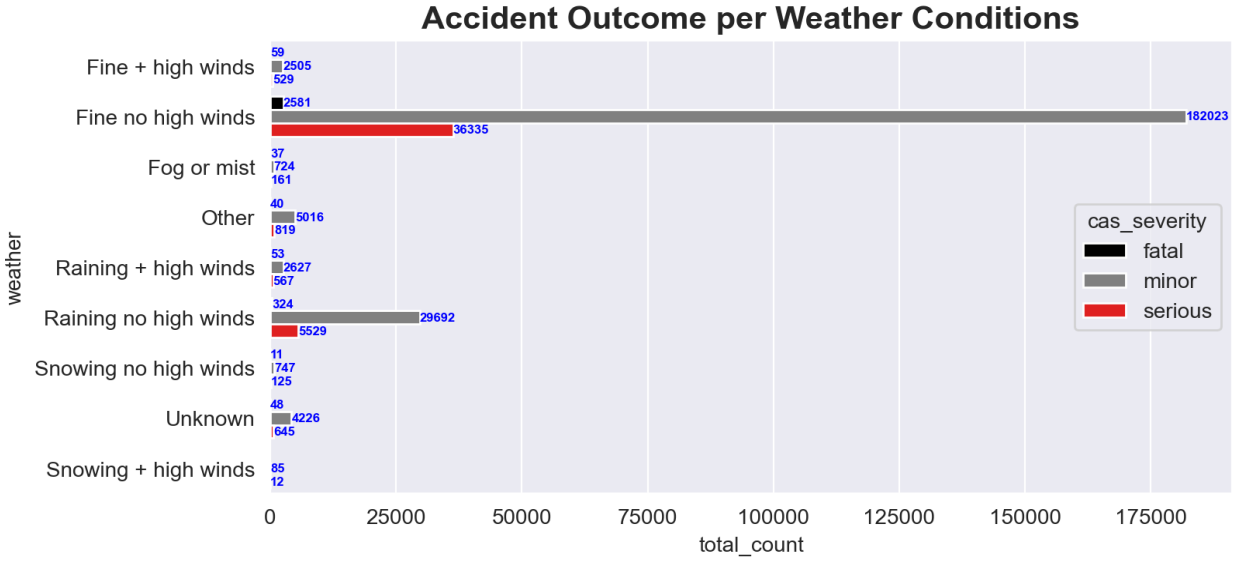

Highest Occurrences of Vehicle Type, Left Hand Drive & Junction Location
junction_location
Mid Junction - on roundabout or on main road
Approaching junction or waiting/parked at junction approach
Not at or within 20 metres of junction
Vehicle_type & was_vehicle_left_hand_drive
vehicle_type: Taxi/Private hire car & was_vehicle_left_hand_drive No
1080
1463
1814
vehicle_type: Motorcycle over 500cc & was_vehicle_left_hand_drive No
1102
1587
2578
vehicle_type: Bus or coach (17 or more pass seats) & was_vehicle_left_hand_drive No
1491
1997
vehicle_type: Motorcycle 125cc and under & was_vehicle_left_hand_drive No
1613
2689
2428
vehicle_type: Van / Goods 3.5 tonnes mgw or under & was_vehicle_left_hand_drive No
2072
3733
7889
vehicle_type: Goods 7.5 tonnes mgw and over & was_vehicle_left_hand_drive No
2875
vehicle_type: Pedal cycle & was_vehicle_left_hand_drive No
4384
4055
4785
vehicle_type: Car & was_vehicle_left_hand_drive No
11.5%
31664
16.9%
46388
33.4%
91919
0
20000
40000
60000
80000
Accidents

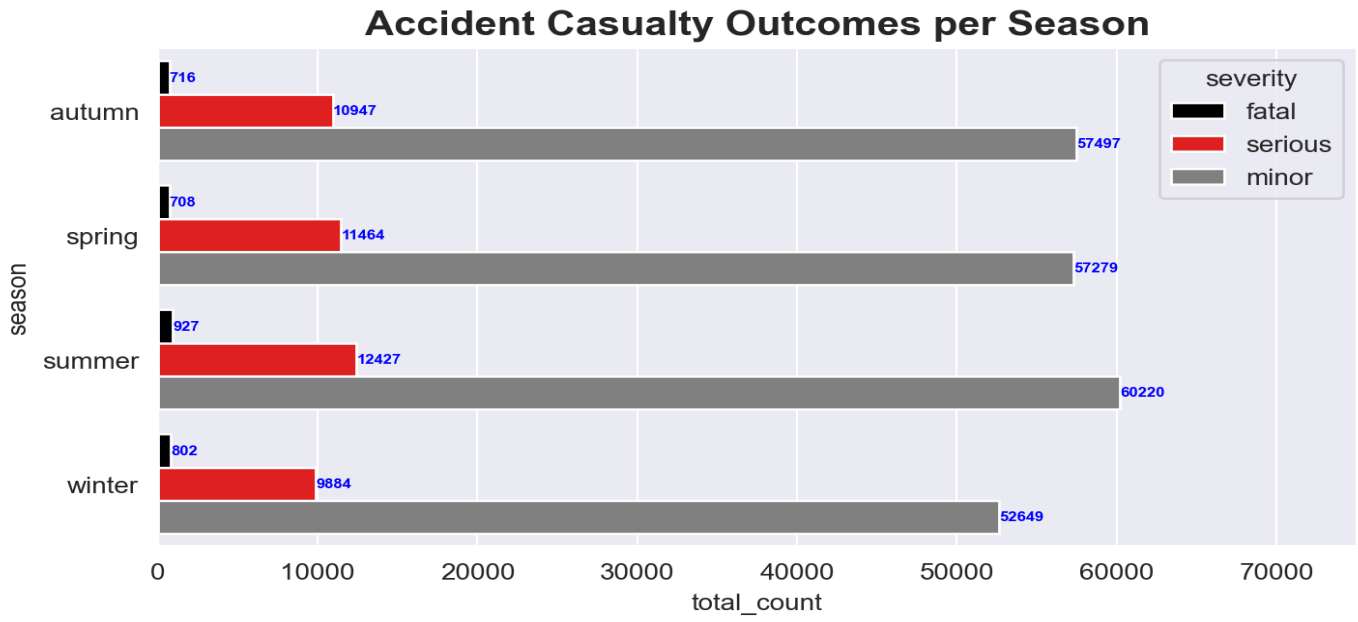

Highest Occurrences of propulsion_code, age_of_vehicle & vehicle_type
Propulsion_code & age_of_vehicle
propulsion_code: Heavy oil & age_of_vehicle 3.0
propulsion_code: Heavy oil & age_of_vehicle 7.0
propulsion_code: Petrol & age_of_vehicle 1.0
propulsion_code: Petrol & age_of_vehicle 9.0
propulsion_code: Undefined & age_of_vehicle 7.0
6265
6266
6951
7597
30954
15523
11.2%
5.6%
vehicle_type
Car
Pedal cycle
0
5000
10000
15000
20000
25000
30000
35000
Accidents

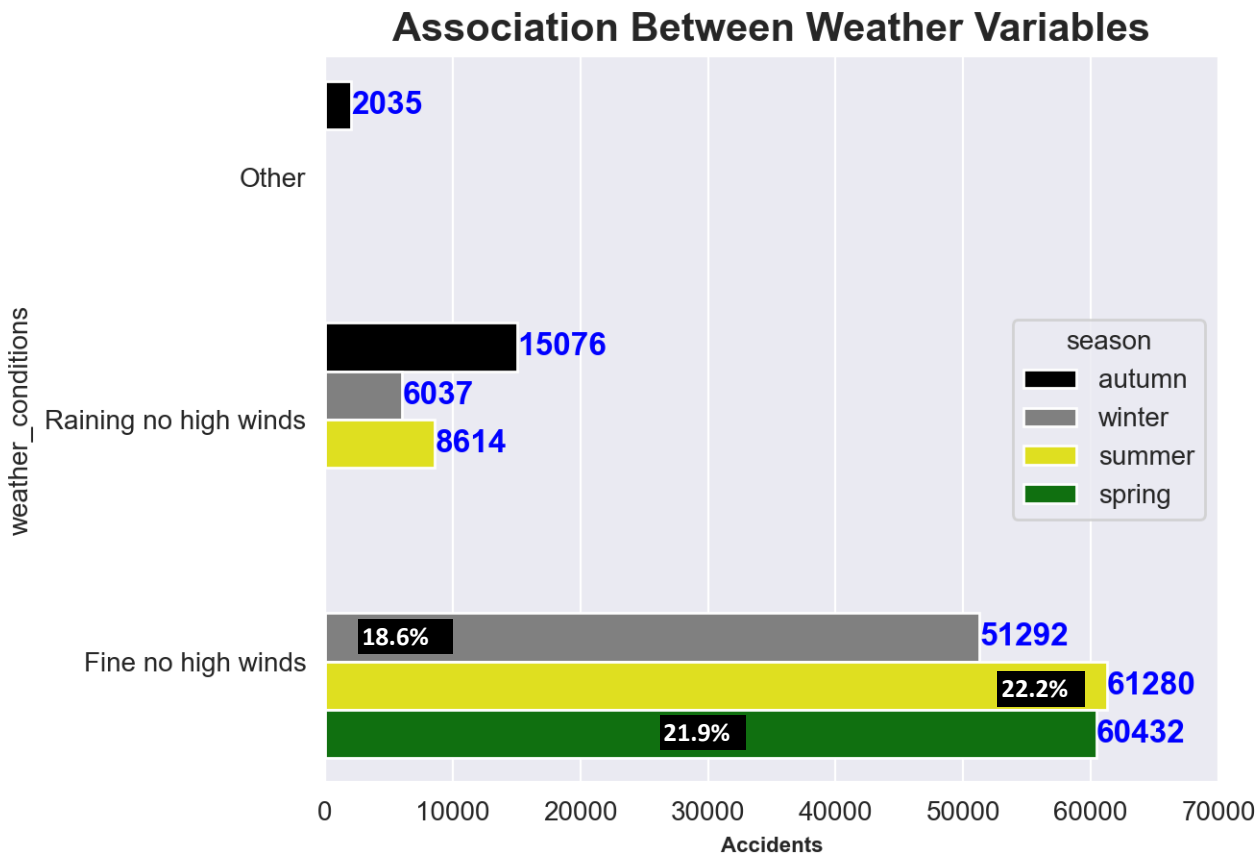

Highest Occurrences of vehicle_type, speed_limit & propulsion_code
Vehicle_type & speed_limit
vehicle_type: Taxi/Private hire car & speed_limit 30 MPH
vehicle_type: Motorcycle over 500cc & speed_limit 30 MPH
vehicle_type: Pedal cycle & speed_limit 20 MPH
vehicle_type: Van / Goods 3.5 tonnes mgw or under & speed_limit 60 MPH
vehicle_type: Bus or coach (17 or more pass seats) & speed_limit 30 MPH
vehicle_type: Motorcycle 125cc and under & speed_limit 30 MPH
vehicle_type: Van / Goods 3.5 tonnes mgw or under & speed_limit 30 MPH
vehicle_type: Pedal cycle & speed_limit 30 MPH
vehicle_type: Car & speed_limit 30 MPH
2044
2185
2214
2364
2650
4286
6152
10998
36030
52464
13.1%
19%
propulsion_code
Heavy oil
Petrol
Undefined
0
10000
20000
30000
40000
50000
60000
70000
Accidents

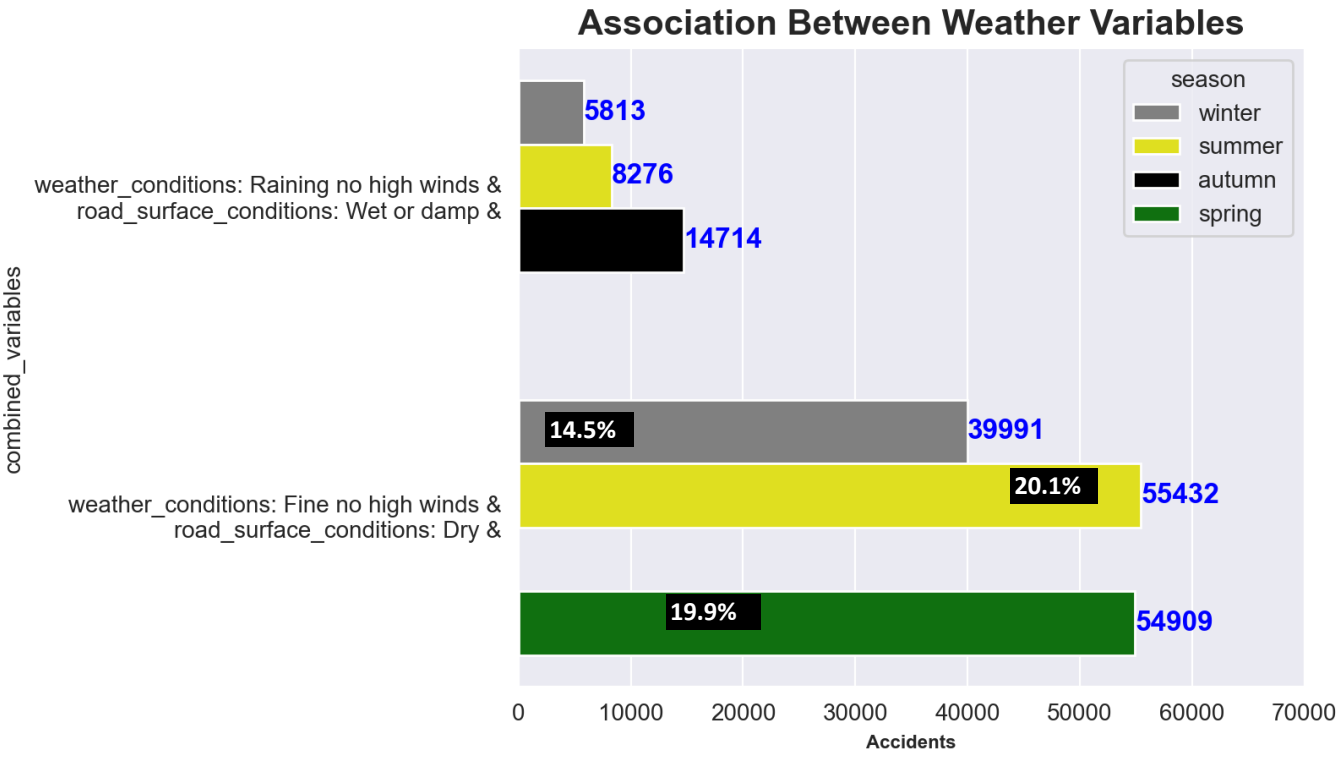
WEATHER INFLUENCE ON ACCIDENTS





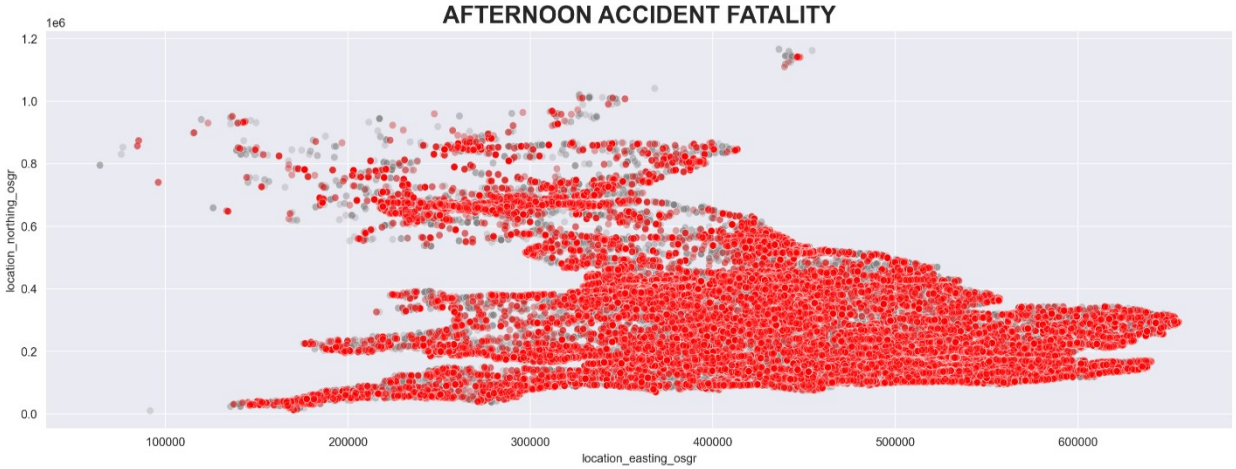


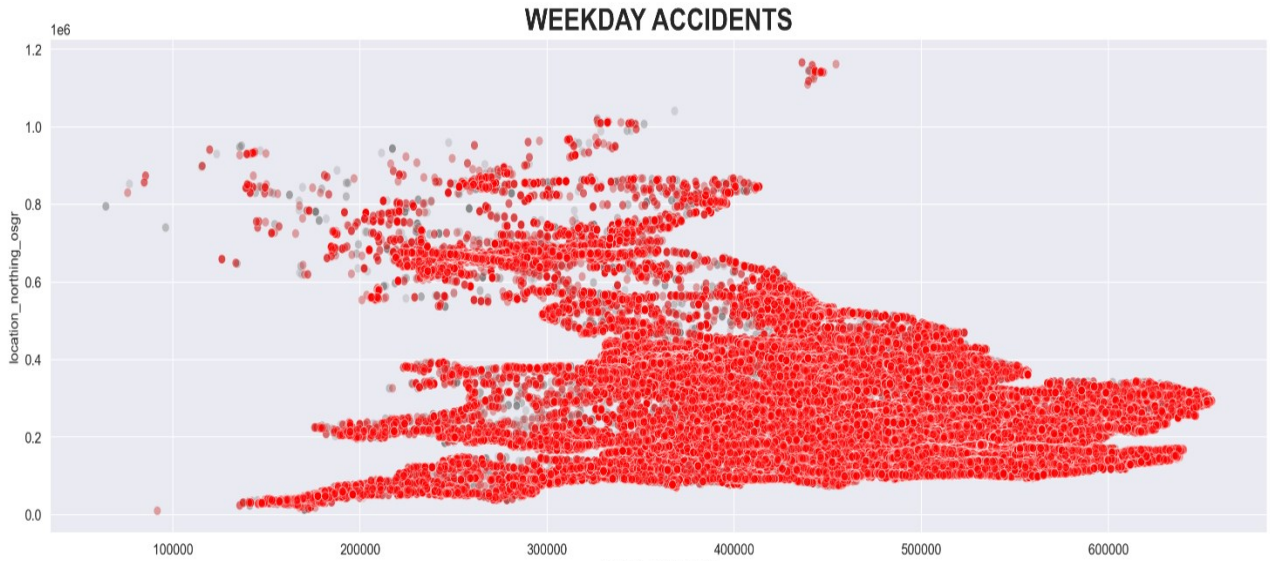

Association Between Weather Variables
weather_conditions
Other
Raining no high winds
Fine no high winds
2035
15076
6037
8614
18.6%
51292
22.2%
61280
21.9%
60432
season
autumn
winter
summer
spring
0
10000
20000
30000
40000
50000
60000
70000
Accidents

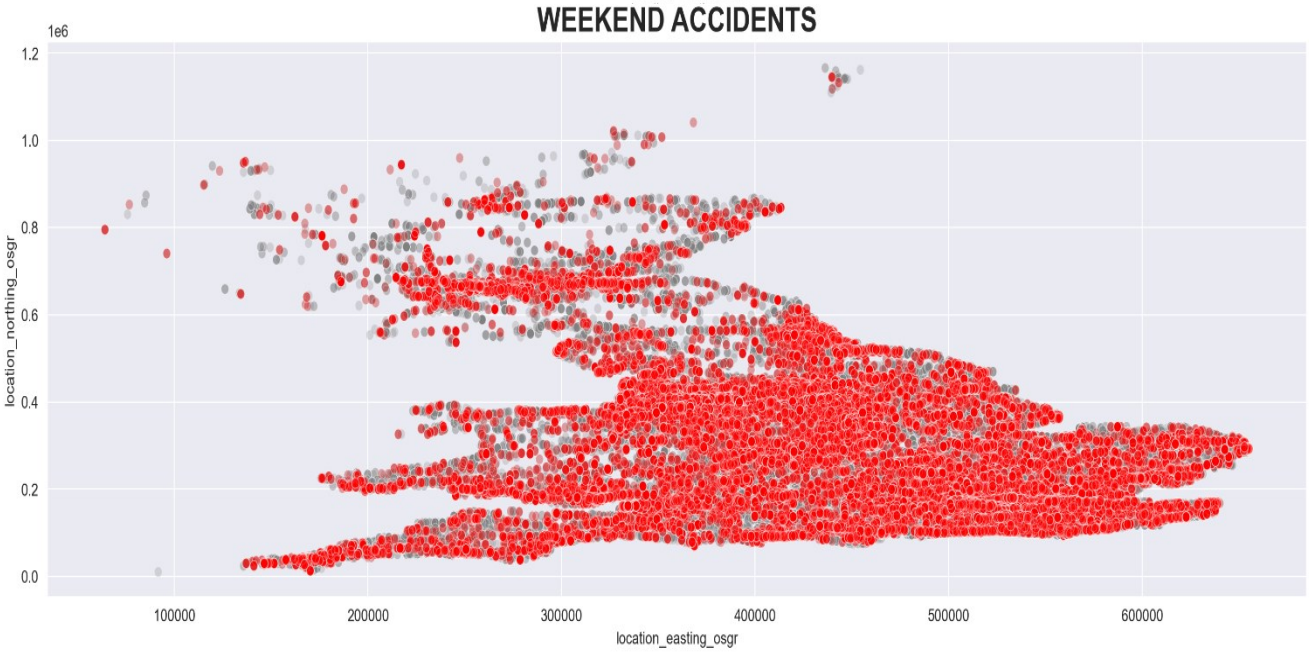

Association Between Weather Variables
combined_variables
weather_conditions: Raining no high winds & road_surface_conditions: Wet or damp &
weather_conditions: Fine no high winds & road_surface_conditions: Dry &
5813
8276
14715
14.5%
39991
20.1%
55432
19.9%
54909
season
winter
summer
autumn
spring
0
10000
20000
30000
40000
50000
60000
70000
Accidents

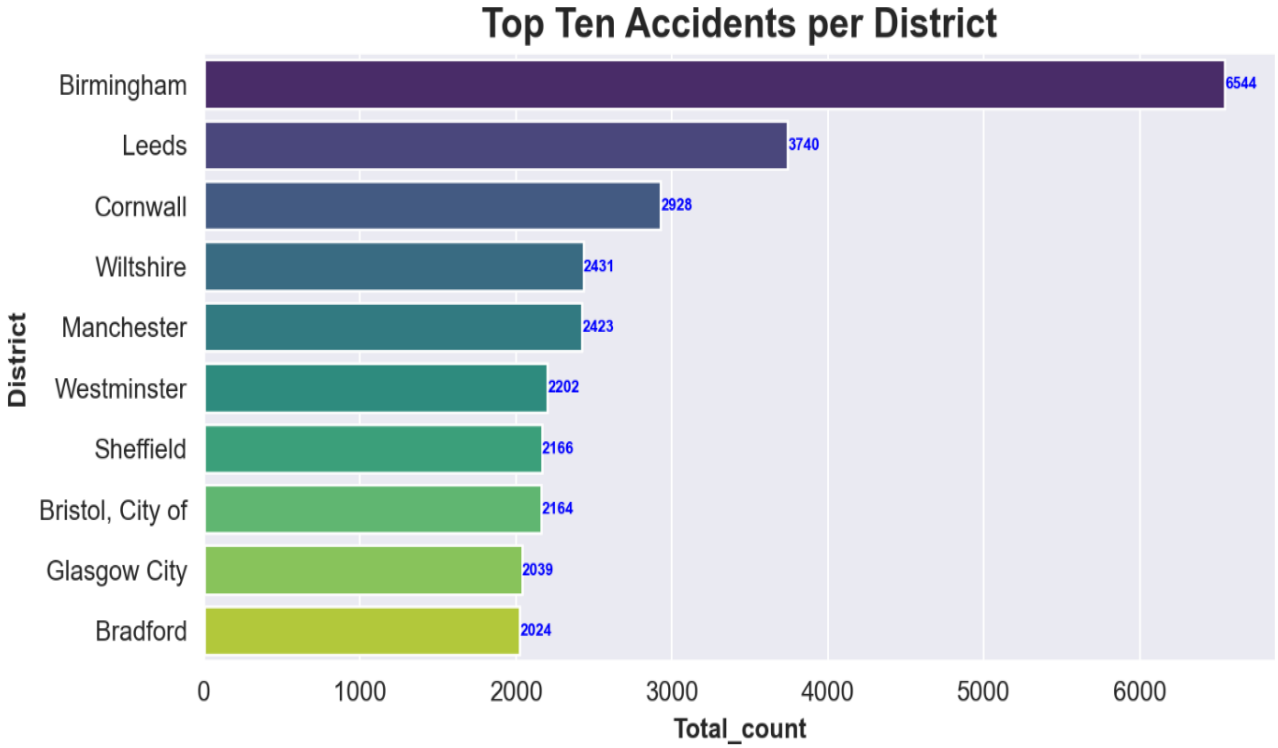
INFLUENCE OF GEOGRAPHIC LOCATION

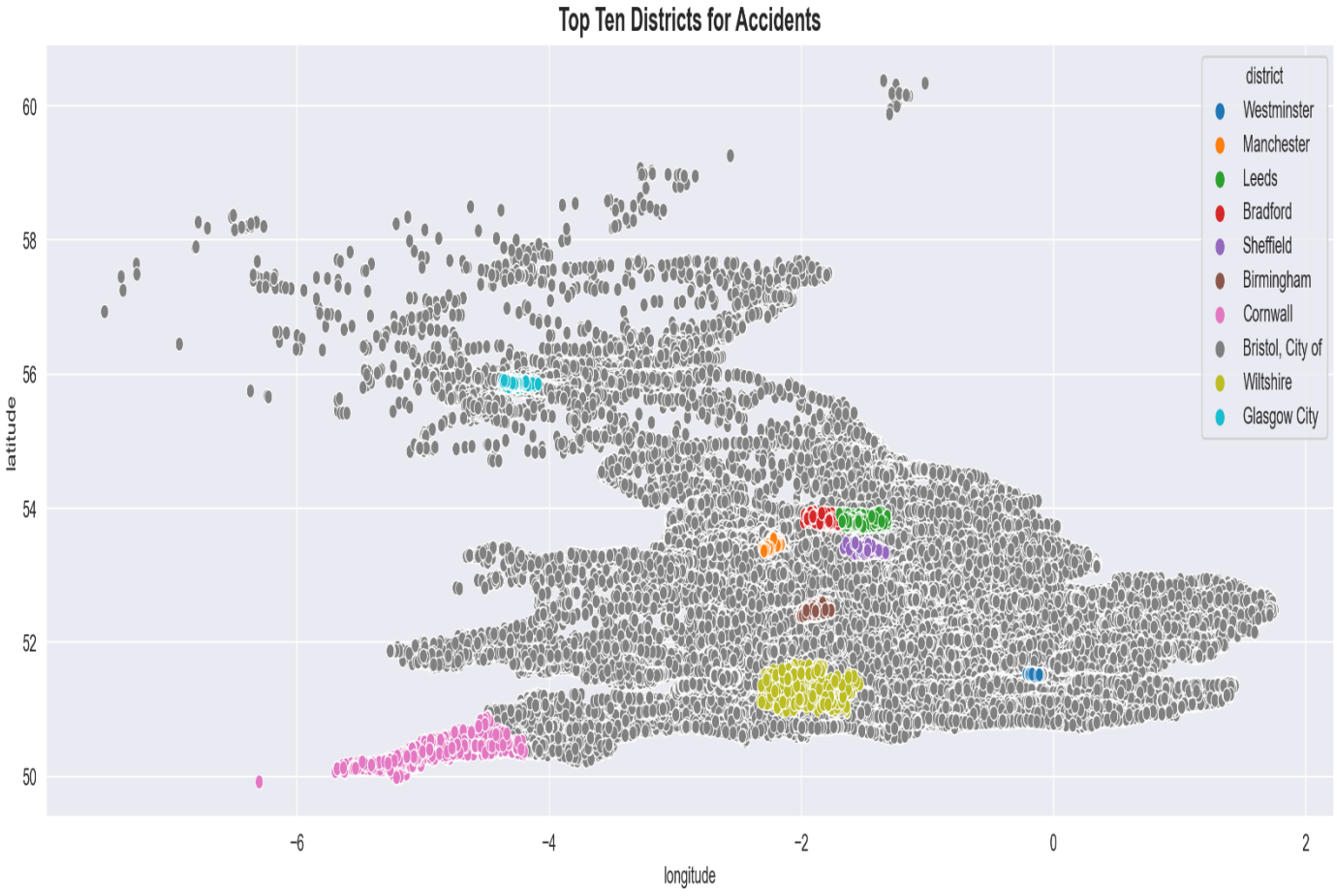
a) CASUALTY OUTCOME ON THE MAP

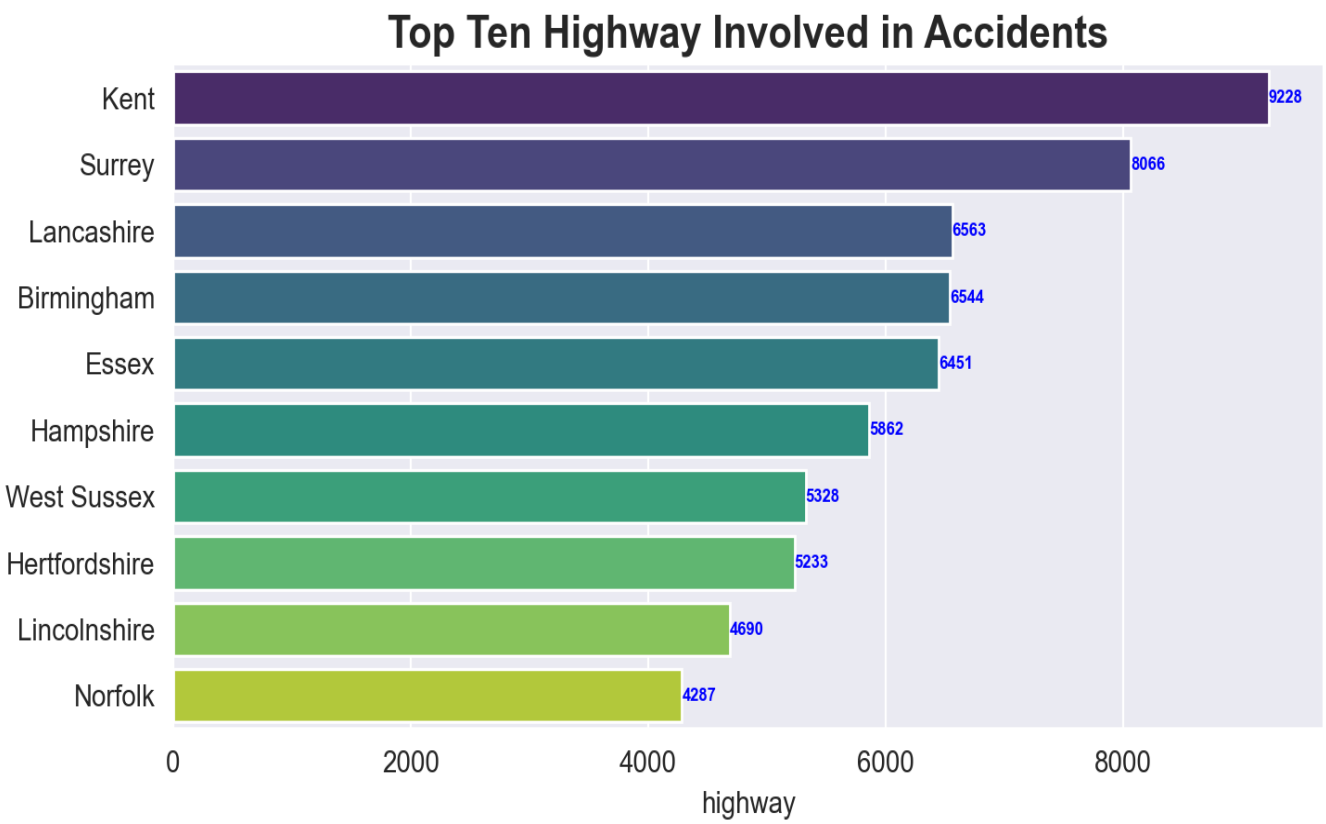


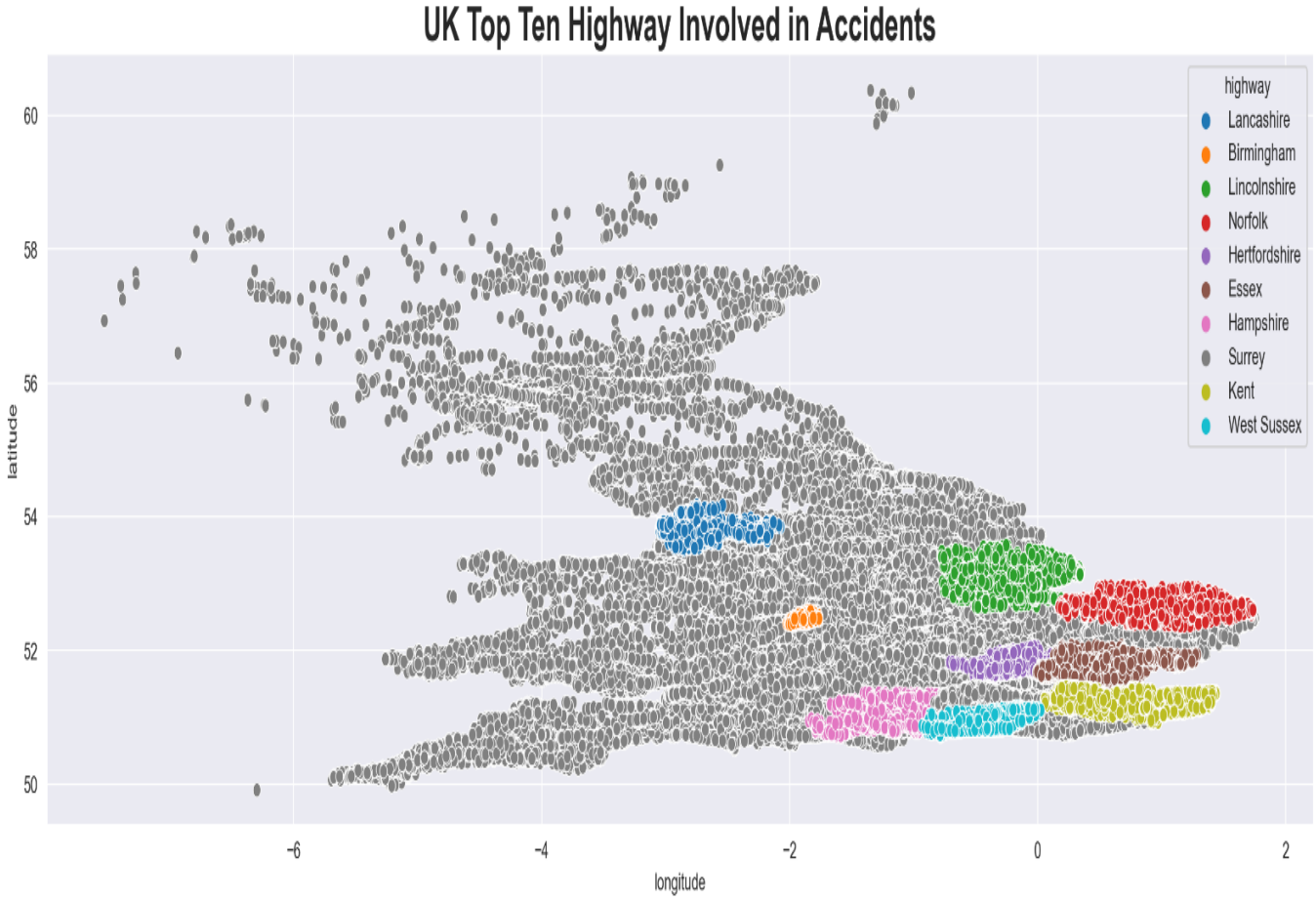

WEEKDAY ACCIDENTS
1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
location_northing_osgr
100000
200000
300000
400000
500000
600000

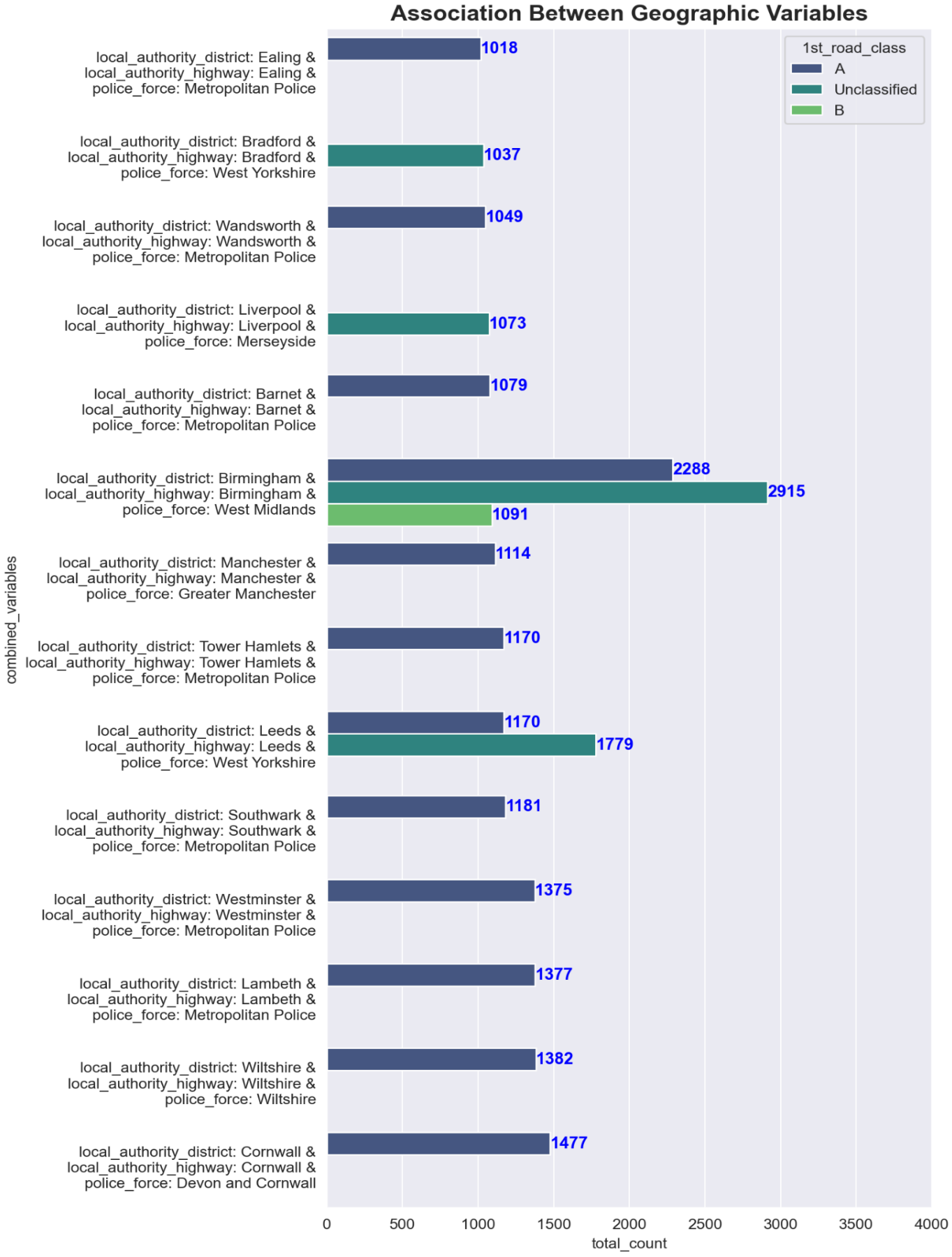

WEEKEND ACCIDENTS
1e6
1.2
1.0
0.8
0.6
0.4
0.2
0.0
location_northing_osgr
100000
200000
300000
400000
500000
600000
location_easting_osgr

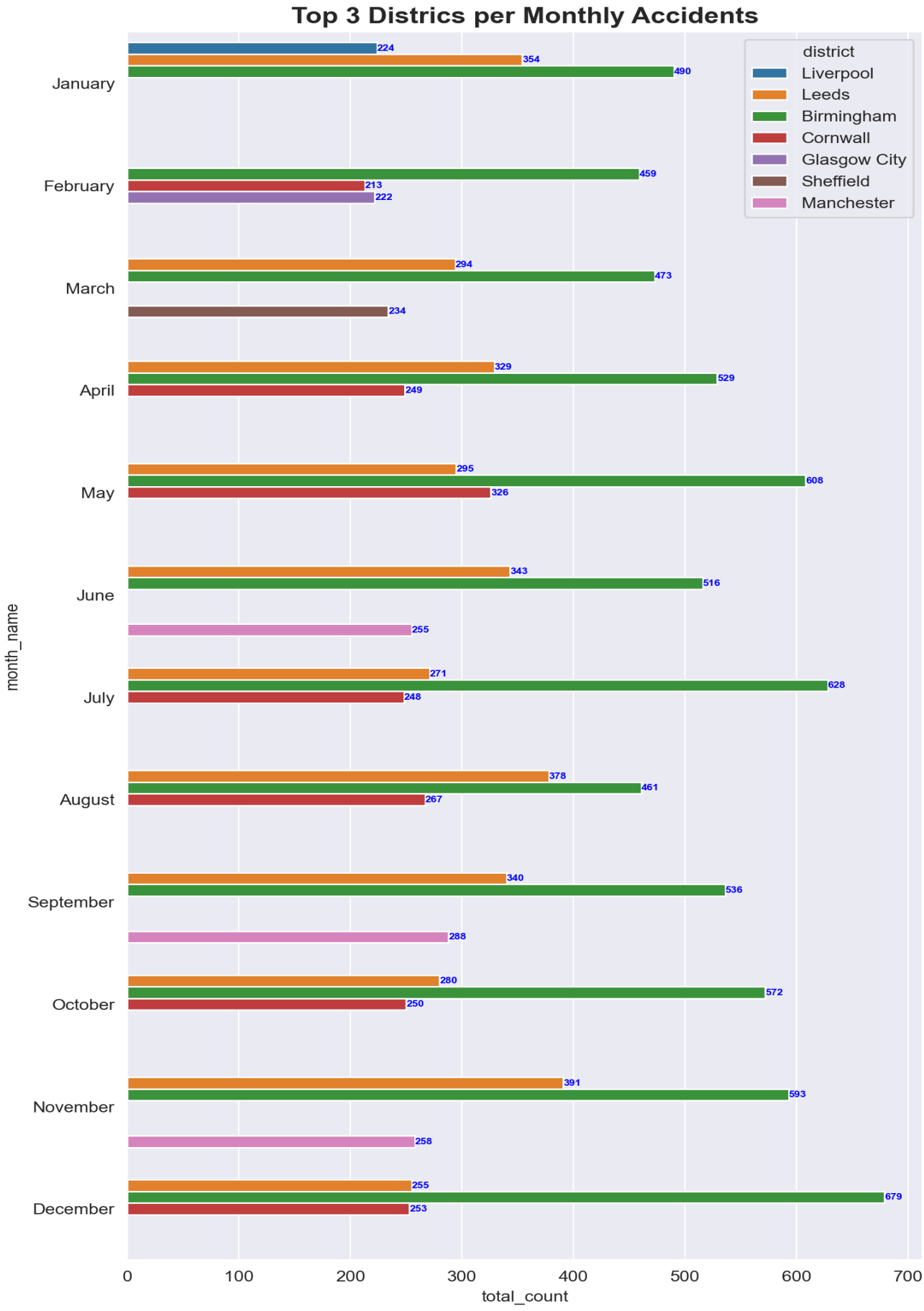

Top Ten Accidents per District
Birmingham
Leeds
Cornwall
Wiltshire
Manchester
Westminster
Sheffield
Bristol, City of
Glasgow City
Bradford
6544
3740
2928
2431
2423
2202
2166
2164
2039
2024
District
0
1000
2000
3000
4000
5000
6000
Total_count

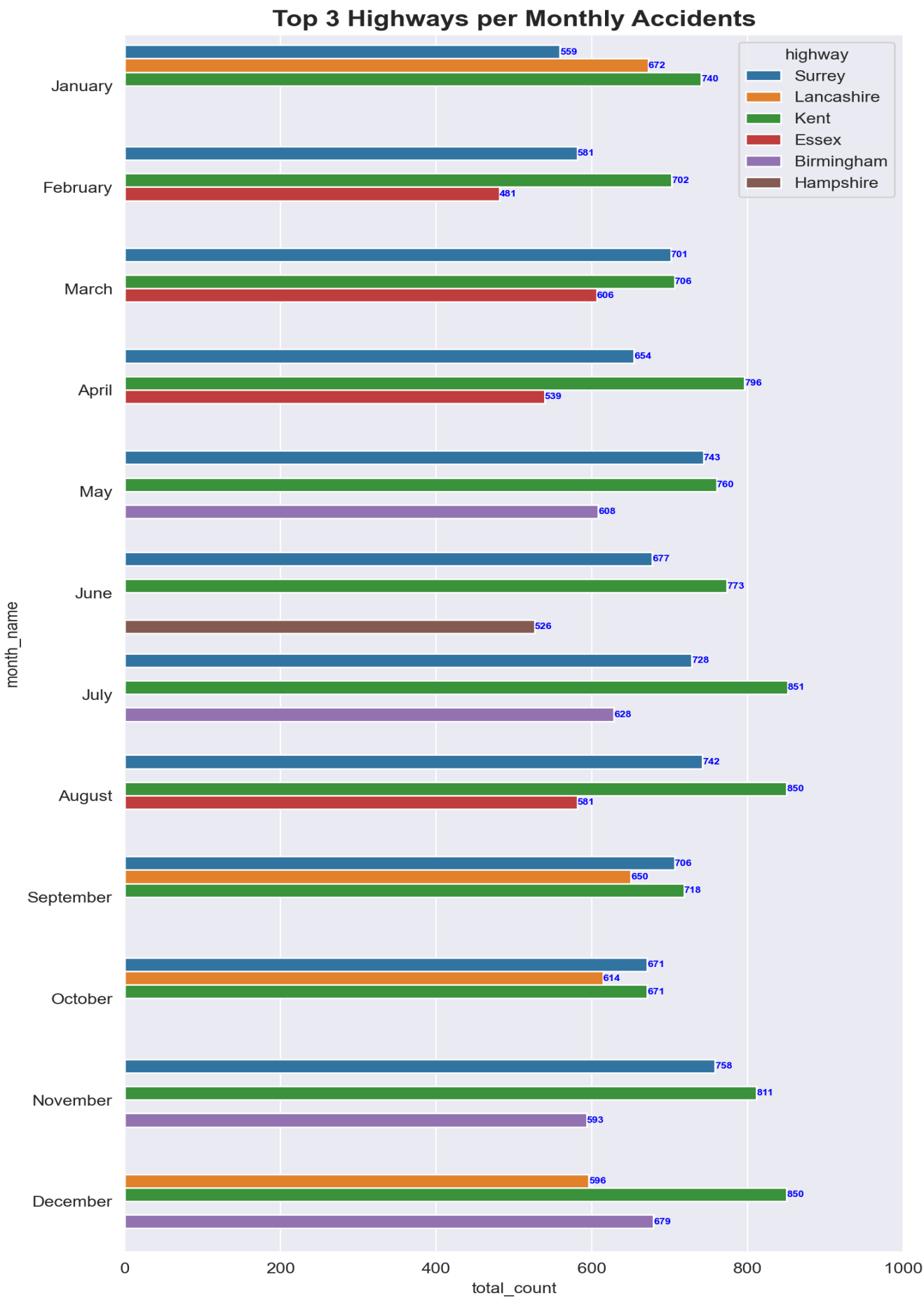

Top Ten Districts for Accidents
district
Westminster
Manchester
Leeds
Bradford
Sheffield
Birmingham
Cornwall
Bristol, City of
Wiltshire
Glasgow City
latitude
60
58
56
54
52
50
-6
-4
-2
0
2
longitude

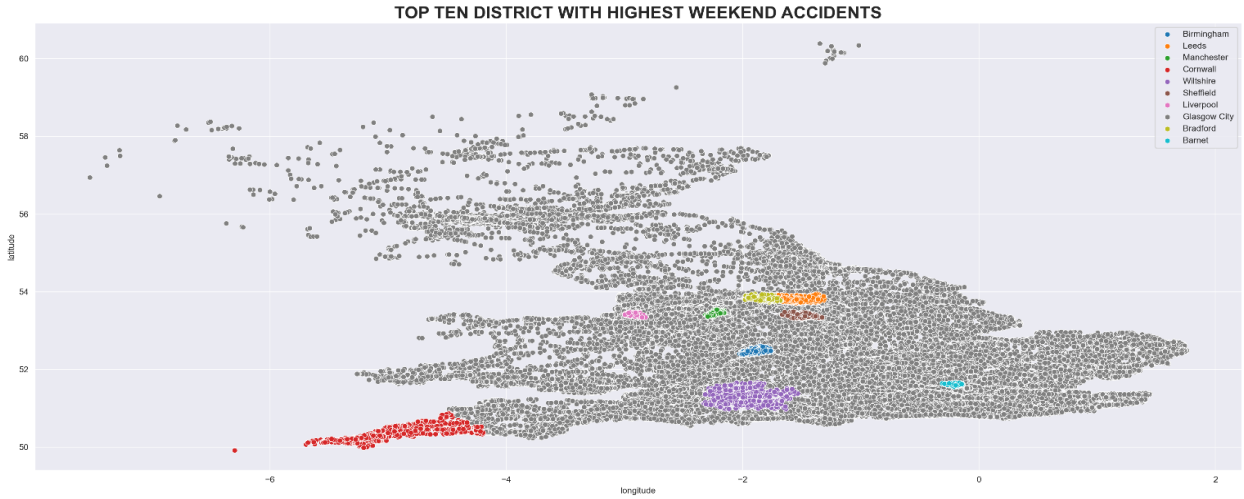

Top Ten Highway Involved in Accidents
Kent
Surrey
Lancashire
Birmingham
Essex
Hampshire
West Sussex
Hertfordshire
Lincolnshire
Norfolk
9228
8066
6563
6544
6451
5862
5328
5233
4690
4287
0
2000
4000
6000
8000
highway

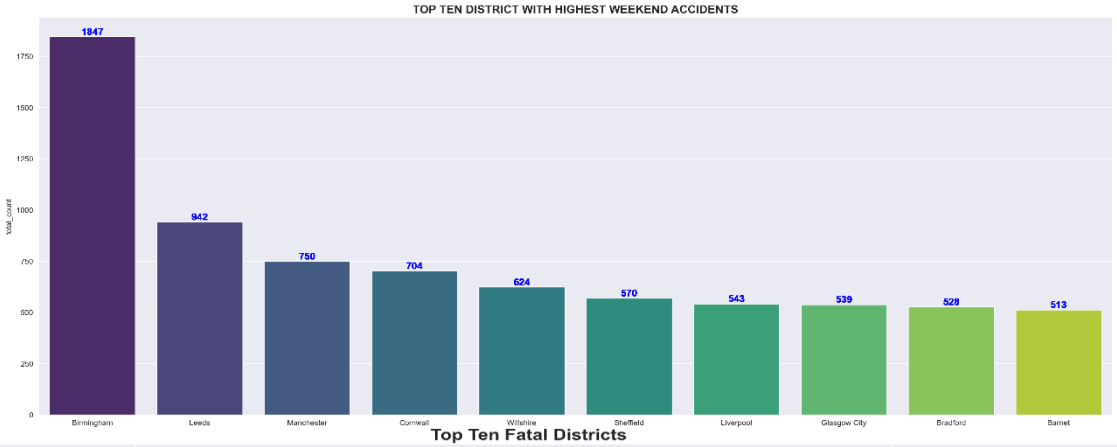

UK Top Ten Highway Involved in Accidents
highway
Lancashire
Birmingham
Lincolnshire
Norfolk
Hertfordshire
Essex
Hampshire
Surrey
Kent
West Sussex
60
58
56
54
52
50
latitude
-6
-4
-2
0
2
longitude

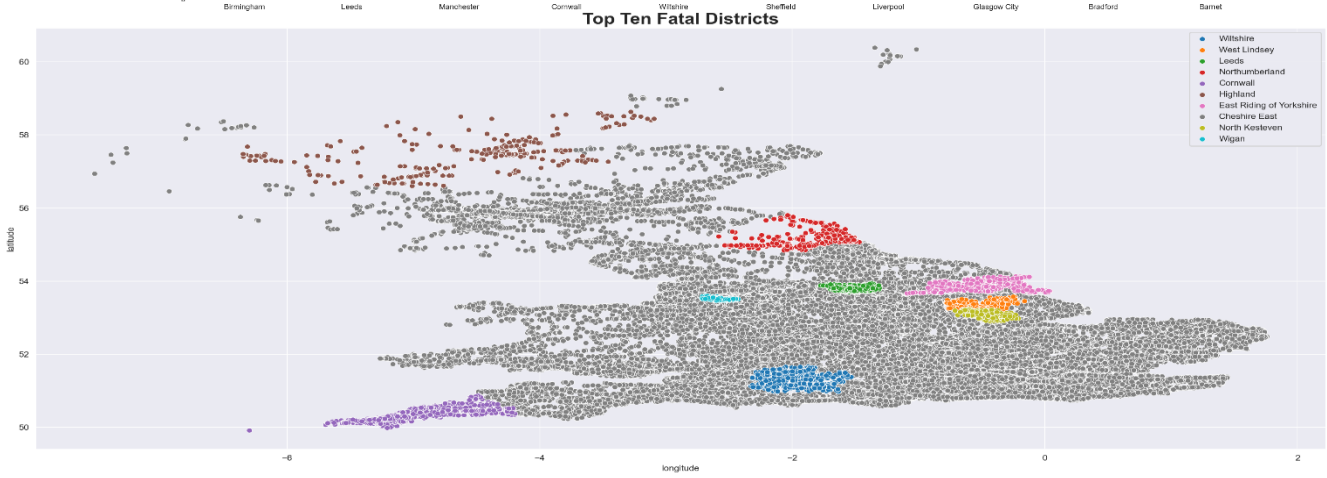

Association Between Geographic Variables
1st_road_class
A
Unclassified
B
local_authority_district: Ealing & local_authority_highway: Ealing & police_force: Metropolitan Police
1018
local_authority_district: Bradford & local_authority_highway: Bradford & police_force: West Yorkshire
1037
local_authority_district: Wandsworth & local_authority_highway: Wandsworth & police_force: Metropolitan Police
1049
local_authority_district: Liverpool & local_authority_highway: Liverpool & police_force: Merseyside
1073
local_authority_district: Barnet & local_authority_highway: Barnet & police_force: Metropolitan Police
1079
local_authority_district: Birmingham & local_authority_highway: Birmingham & police_force: West Midlands
2288
2915
1091
local_authority_district: Manchester & local_authority_highway: Manchester & police_force: Greater Manchester
1114
local_authority_district: Tower Hamlets & local_authority_highway: Tower Hamlets & police_force: Metropolitan Police
1170
local_authority_district: Leeds & local_authority_highway: Leeds & police_force: West Yorkshire
1170
1779
local_authority_district: Southwark & local_authority_highway: Southwark & police_force: Metropolitan Police
1181
local_authority_district: Westminster & local_authority_highway: Westminster & police_force: Metropolitan Police
1375
local_authority_district: Lambeth & local_authority_highway: Lambeth & police_force: Metropolitan Police
1377
local_authority_district: Wiltshire & local_authority_highway: Wiltshire & police_force: Wiltshire
1382
local_authority_district: Cornwall & local_authority_highway: Cornwall & police_force: Devon and Cornwall
1477
combined_variables
0
500
1000
1500
2000
2500
3000
3500
4000
total_count

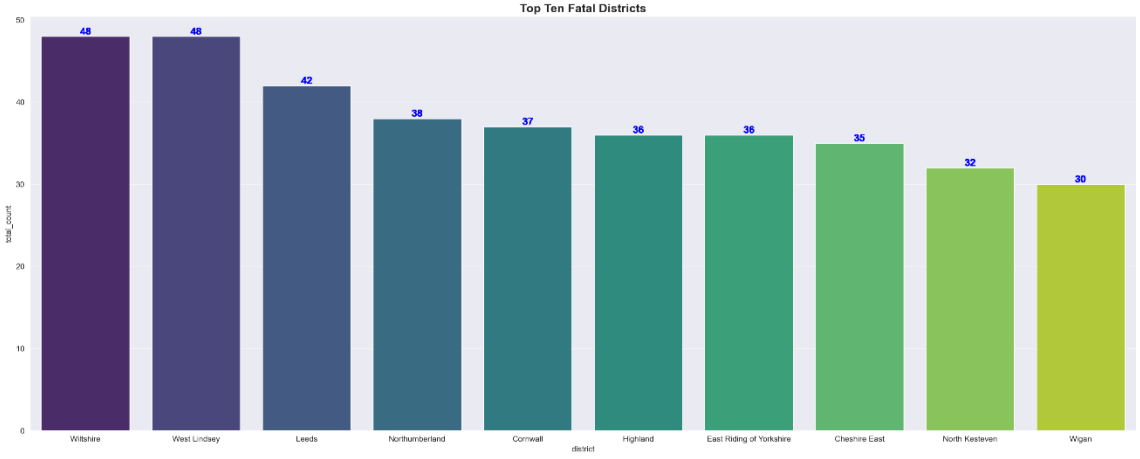

Top 3 Districs per Monthly Accidents
district
Liverpool
Leeds
Birmingham
Cornwall
Glasgow City
Sheffield
Manchester
month_name
January
February
March
April
May
June
July
August
September
October
November
December
224
354
490
459
213
222
294
473
234
329
529
249
295
608
326
343
516
255
271
628
248
378
461
267
340
536
288
280
572
250
391
593
258
255
679
253
0
100
200
300
400
500
600
700
total_count


Top 3 Highways per Monthly Accidents
highway
Surrey
Lancashire
Kent
Essex
Birmingham
Hampshire
January
February
March
April
May
June
July
August
September
October
November
December
559
672
740
581
702
481
701
706
606
654
796
539
743
760
608
677
773
526
728
851
628
742
850
581
706
650
718
671
614
671
758
811
593
596
850
679
month_name
0
200
400
600
800
1000
total_count

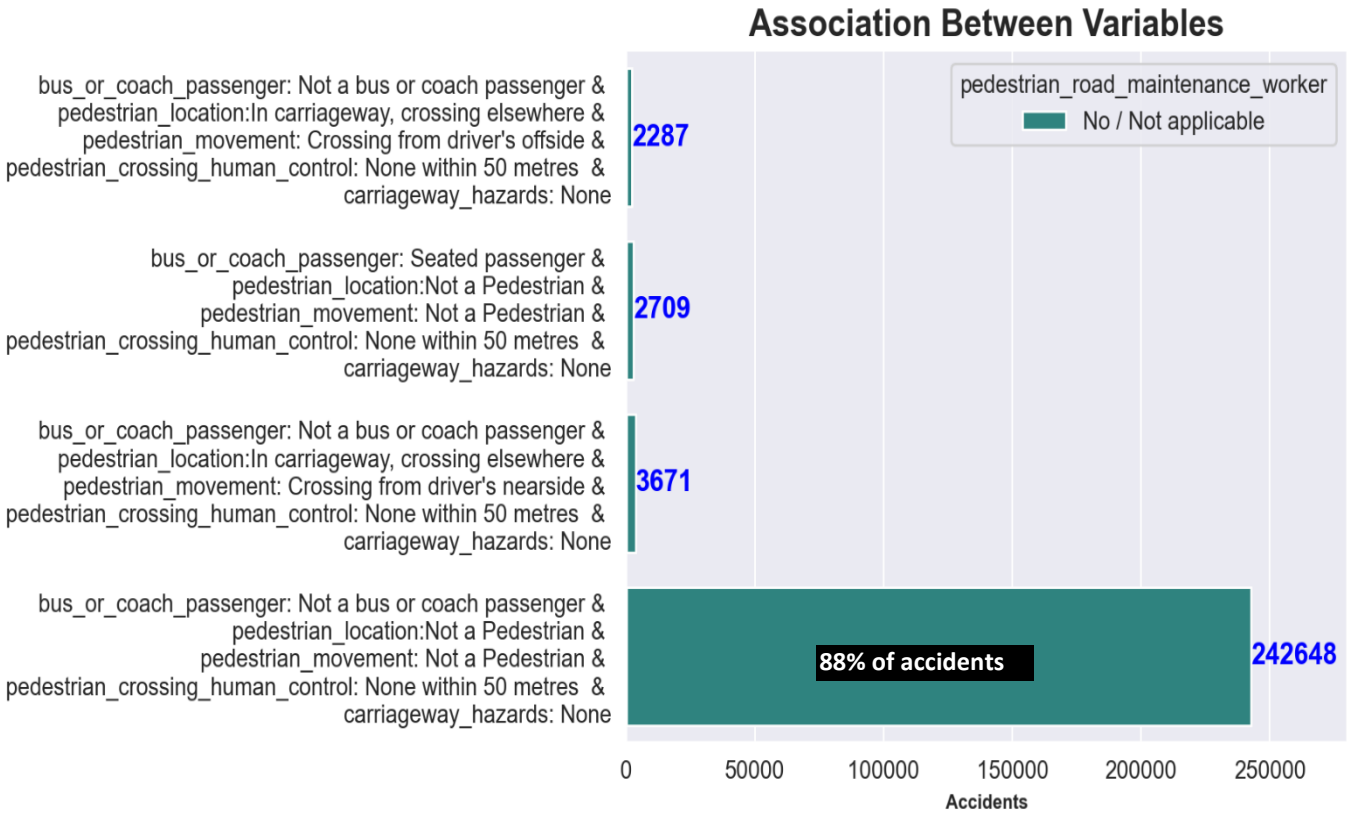

TOP TEN DISTRICT WITH HIGHEST WEEKEND ACCIDENTS
Birmingham
Leeds
Manchester
Cornwall
Wiltshire
Sheffield
Liverpool
Glasgow City
Bradford
Barnet
latitude
longitude

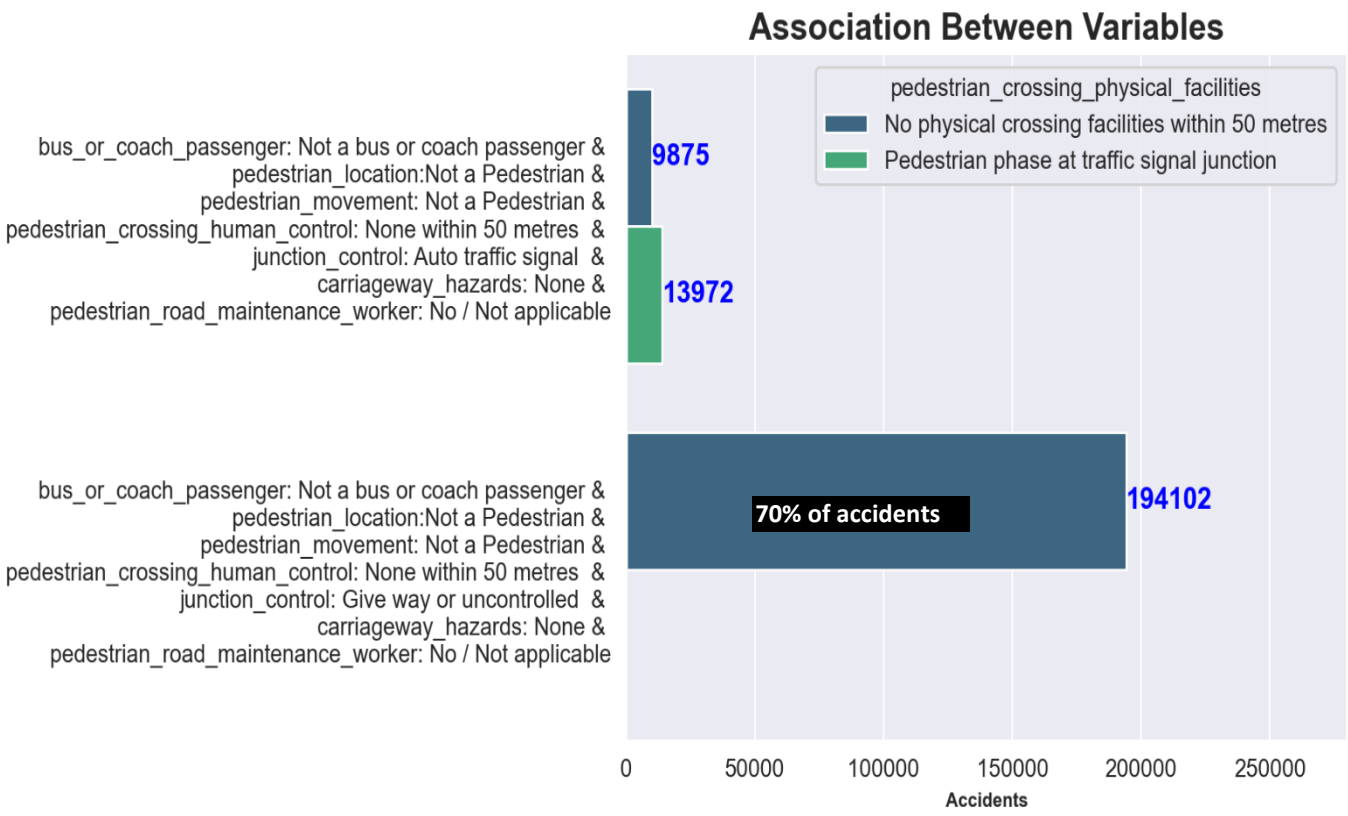

TOP TEN DISTRICT WITH HIGHEST WEEKEND ACCIDENTS
total_count
1847
942
750
704
624
570
543
539
528
513
Birmingham
Leeds
Manchester
Cornwall
Wiltshire
Sheffield
Liverpool
Glasgow City
Bradford
Barnet

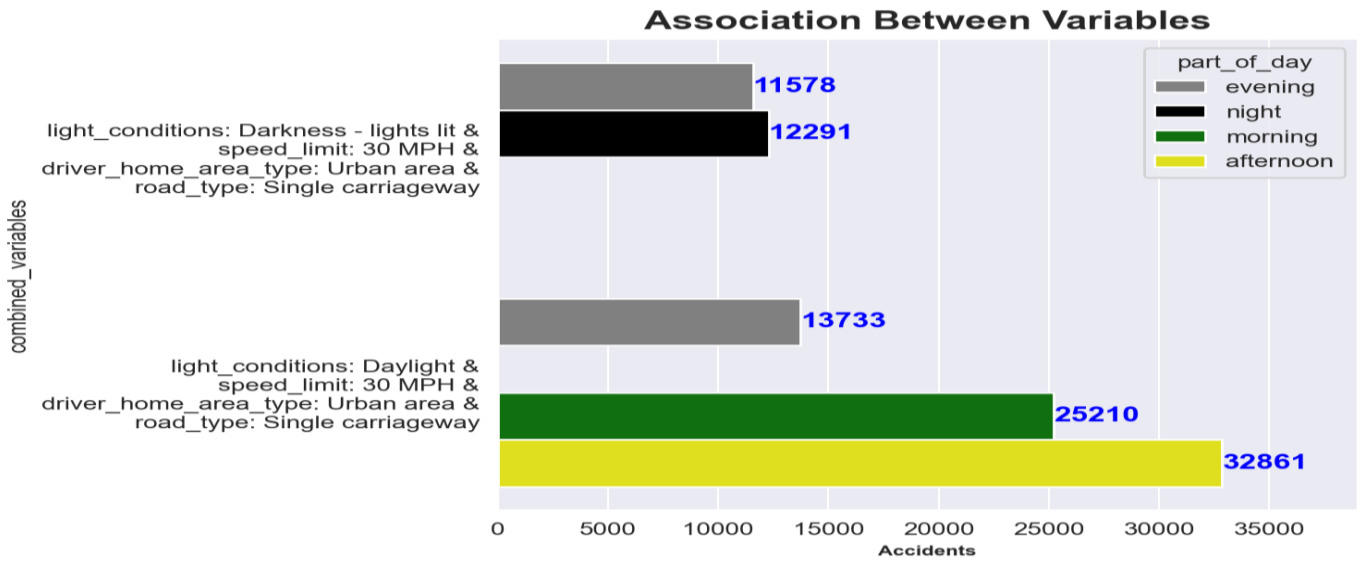

Top Ten Fatal Districts
Wiltshire
West Lindsey
Leeds
Northumberland
Cornwall
Highland
East Riding of Yorkshire
Cheshire East
North Kesteven
Wigan
latitude
longitude

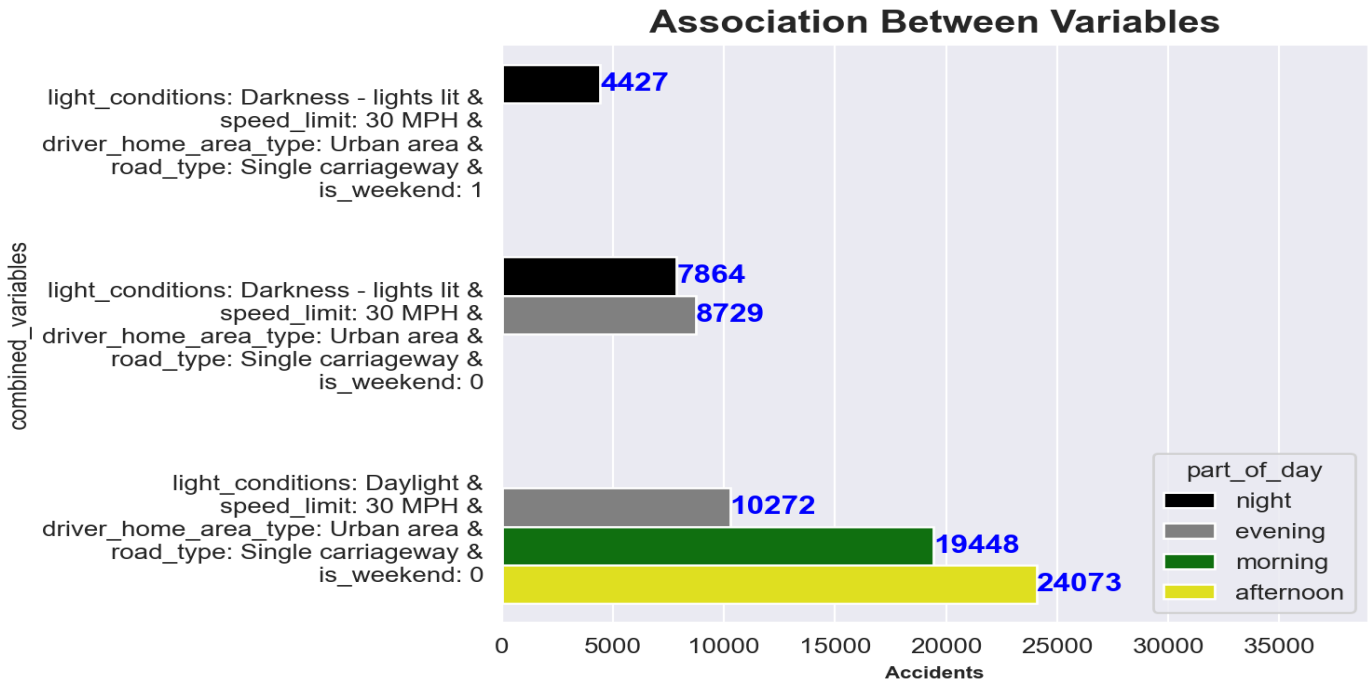

Top Ten Fatal Districts
total_count
48
48
42
38
37
36
36
35
32
30
Wiltshire
West Lindsey
Leeds
Northumberland
Cornwall
Highland
East Riding of Yorkshire
Cheshire East
North Kesteven
Wigan
district

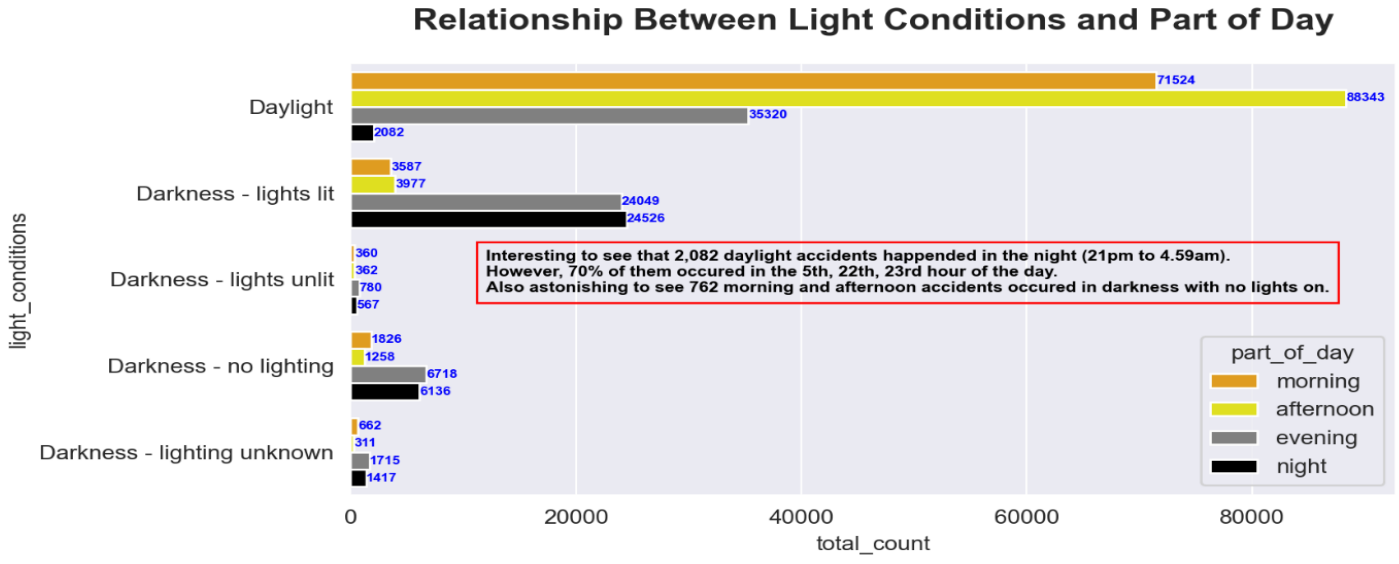
SITUATIONAL INFLUENCE ON ACCIDENTS & OUTCOME

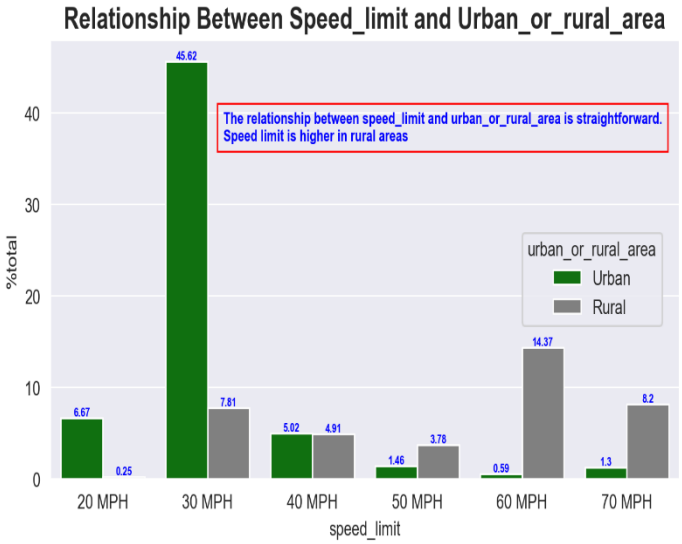


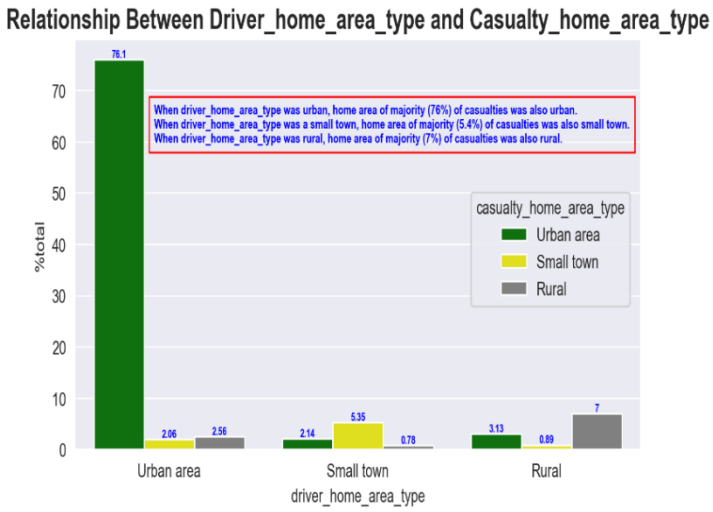


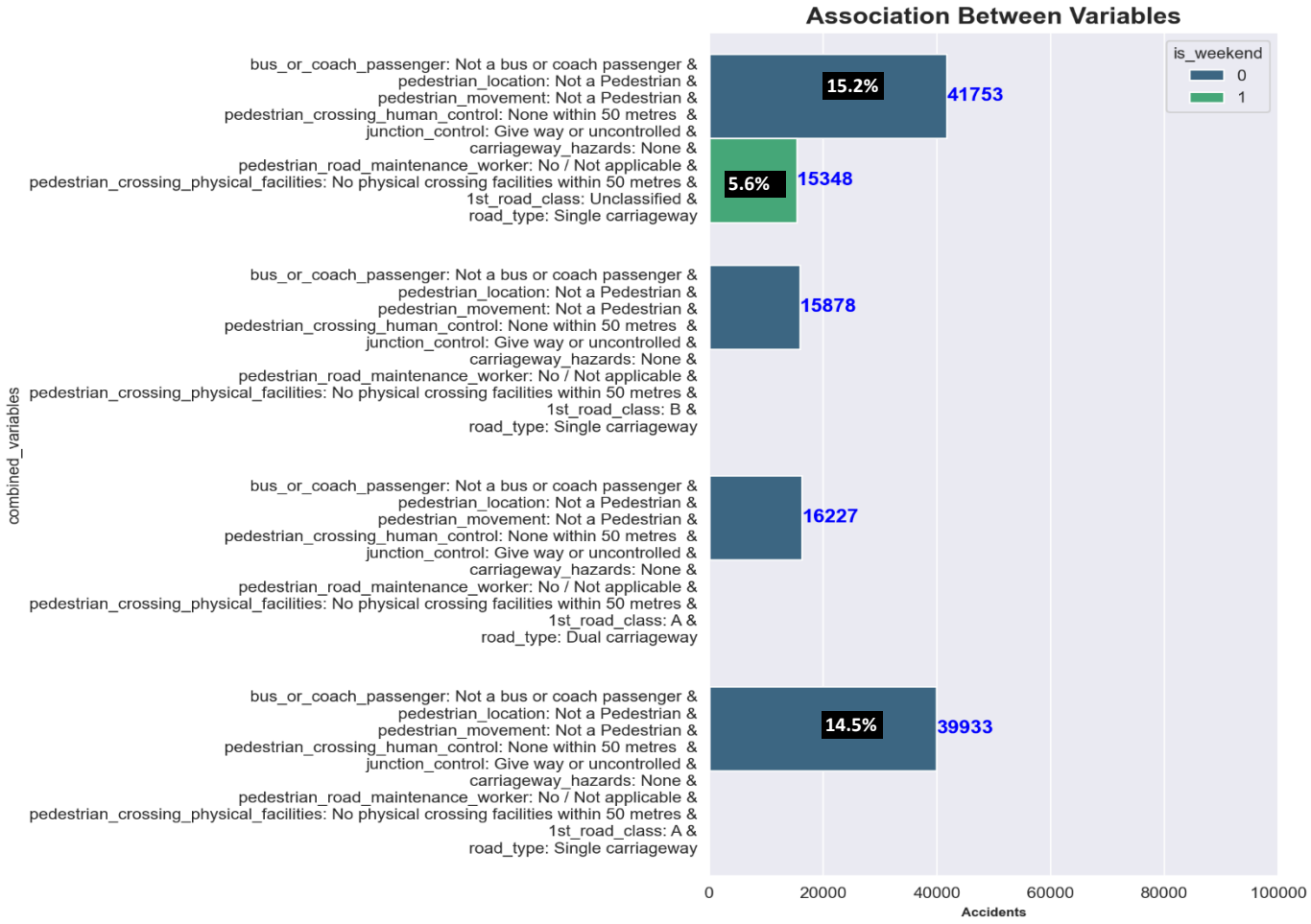

Association Between Variables
part_of_day
evening
night
morning
afternoon
11578
12291
13733
25210
32861
light_conditions: Darkness - lights lit & speed_limit: 30 MPH & driver_home_area_type: Urban area & road_type: Single carriageway
light_conditions: Daylight & speed_limit: 30 MPH & driver_home_area_type: Urban area & road_type: Single carriageway
combined_variables
0
5000
10000
15000
20000
25000
30000
35000
Accidents

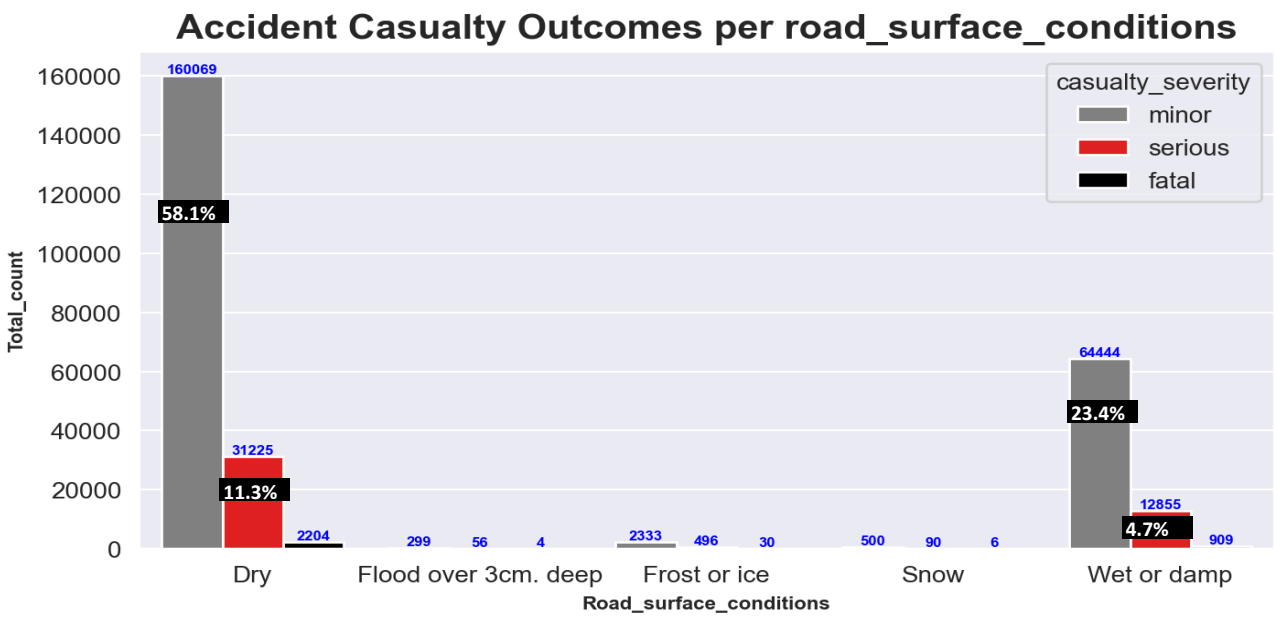

Association Between Variables
4427
7864
8729
10272
19448
24073
light_conditions: Darkness - lights lit & speed_limit: 30 MPH & driver_home_area_type: Urban area & road_type: Single carriageway & is_weekend: 1
light_conditions: Darkness - lights lit & speed_limit: 30 MPH & driver_home_area_type: Urban area & road_type: Single carriageway & is_weekend: 0
light_conditions: Daylight & speed_limit: 30 MPH & driver_home_area_type: Urban area & road_type: Single carriageway & is_weekend: 0
combined_variables
part_of_day
night
evening
morning
afternoon
0
5000
10000
15000
20000
25000
30000
35000
Accidents

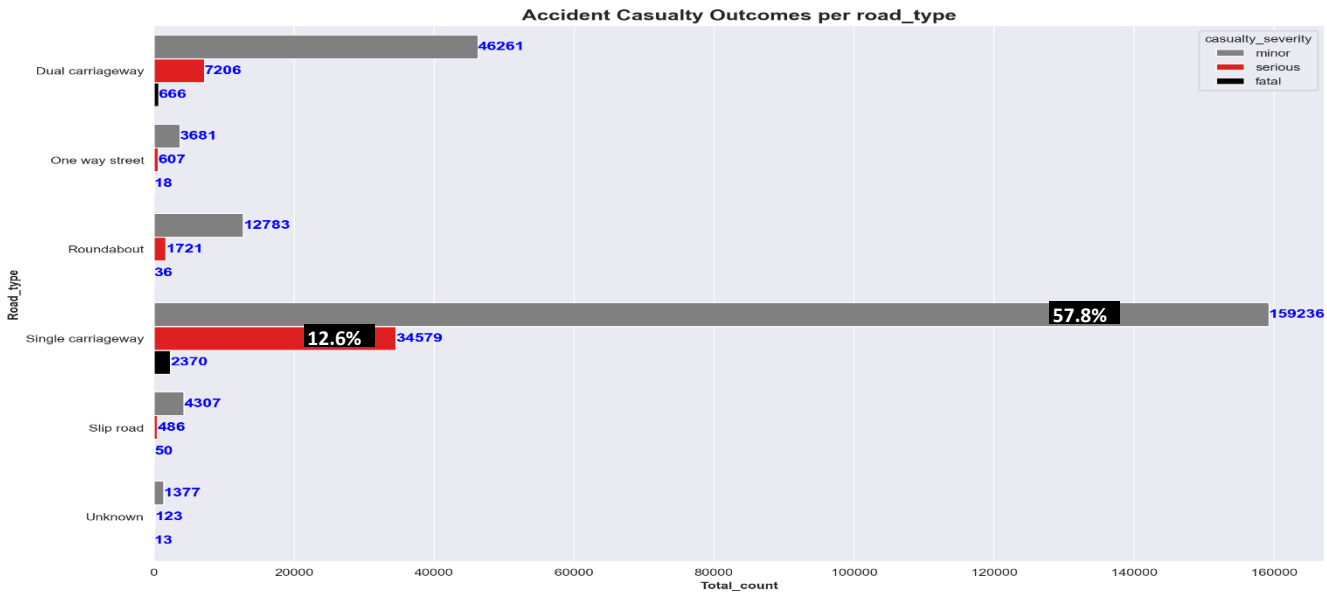

Relationship Between Light Conditions and Part of Day
Daylight
Darkness - lights lit
Darkness - lights unlit
Darkness - no lighting
Darkness - lighting unknown
light_conditions
71524
88343
35320
2082
3587
3977
24049
24526
360
362
780
567
1826
1258
6718
6136
662
311
1715
1417
Interesting to see that 2,082 daylight accidents happended in the night (21pm to 4.59am).
However, 70% of them occured in the 5th, 22th, 23rd hour of the day.
Also astonishing to see 762 morning and afternoon accidents occured in darkness with no lights on.
part_of_day
morning
afternoon
evening
night
0
20000
40000
60000
80000
total_count

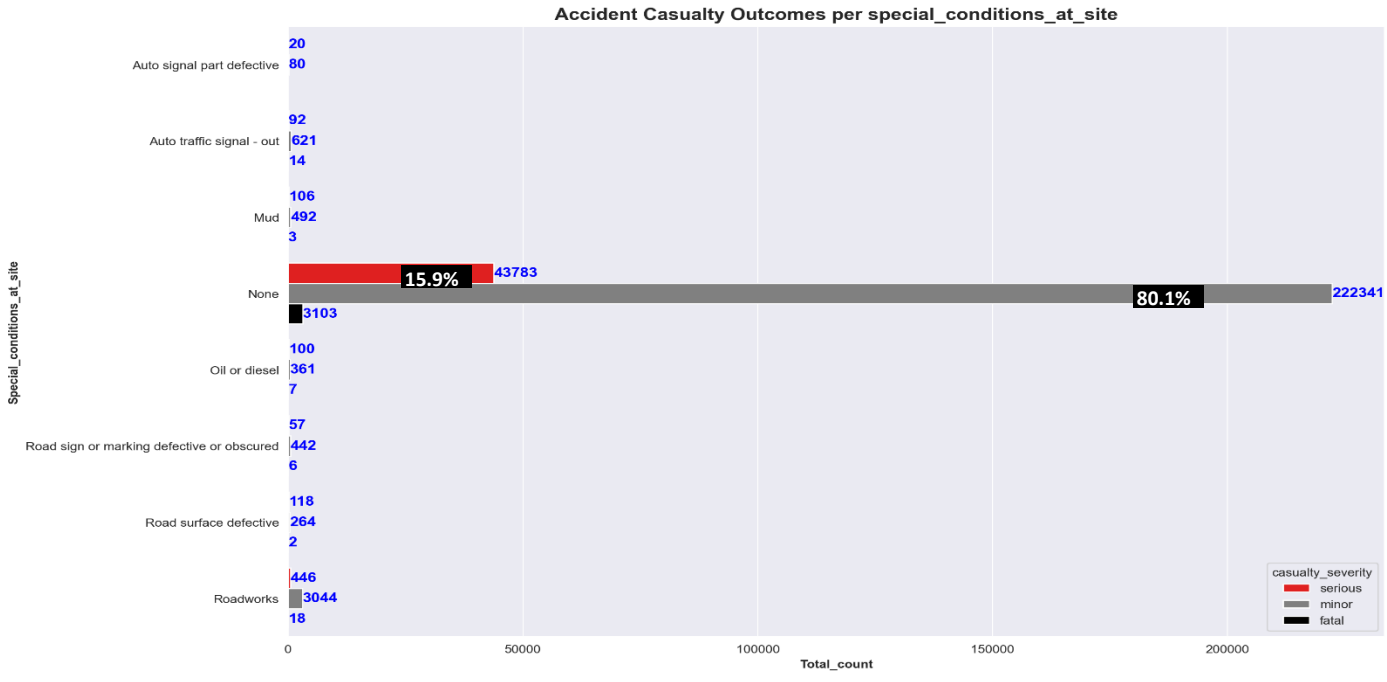

Relationship Between Speed_limit and Urban_or_rural_area
The relationship between speed_limit and urban_or_rural_area is straightforward.
Speed limit is higher in rural areas
urban_or_rural_area
Urban
Rural
%total
speed_limit
20 MPH
30 MPH
40 MPH
50 MPH
60 MPH
70 MPH
6.67
0.25
45.62
7.81
5.02
4.91
1.46
3.78
0.59
14.37
1.3
8.2

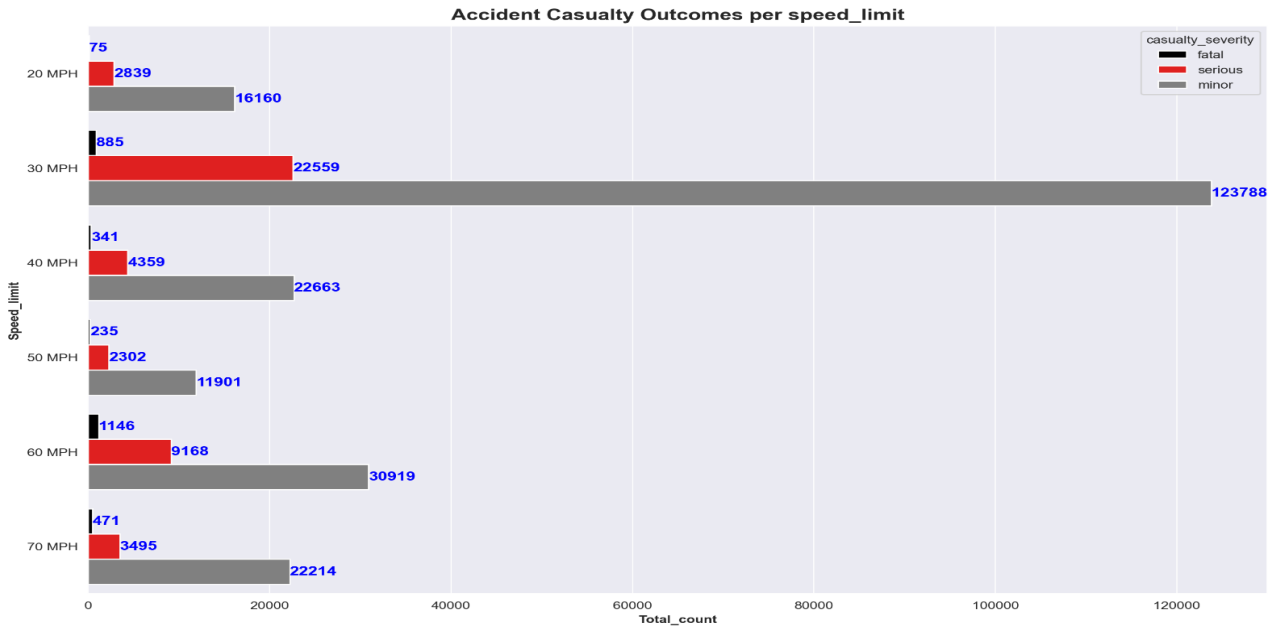

Relationship Between Driver_home_area_type and Casualty_home_area_type
When driver_home_area_type was urban, home area of majority (76%) of casualties was also urban.
When driver_home_area_type was a small town, home area of majority (5.4%) of casualties was also small town.
When driver_home_area_type was rural, home area of majority (7%) of casualties was also rural.
casualty_home_area_type
Urban area
Small town
Rural
%total
driver_home_area_type
Urban area
Small town
Rural
76.1
2.06
2.56
2.14
5.35
0.78
3.13
0.89
7

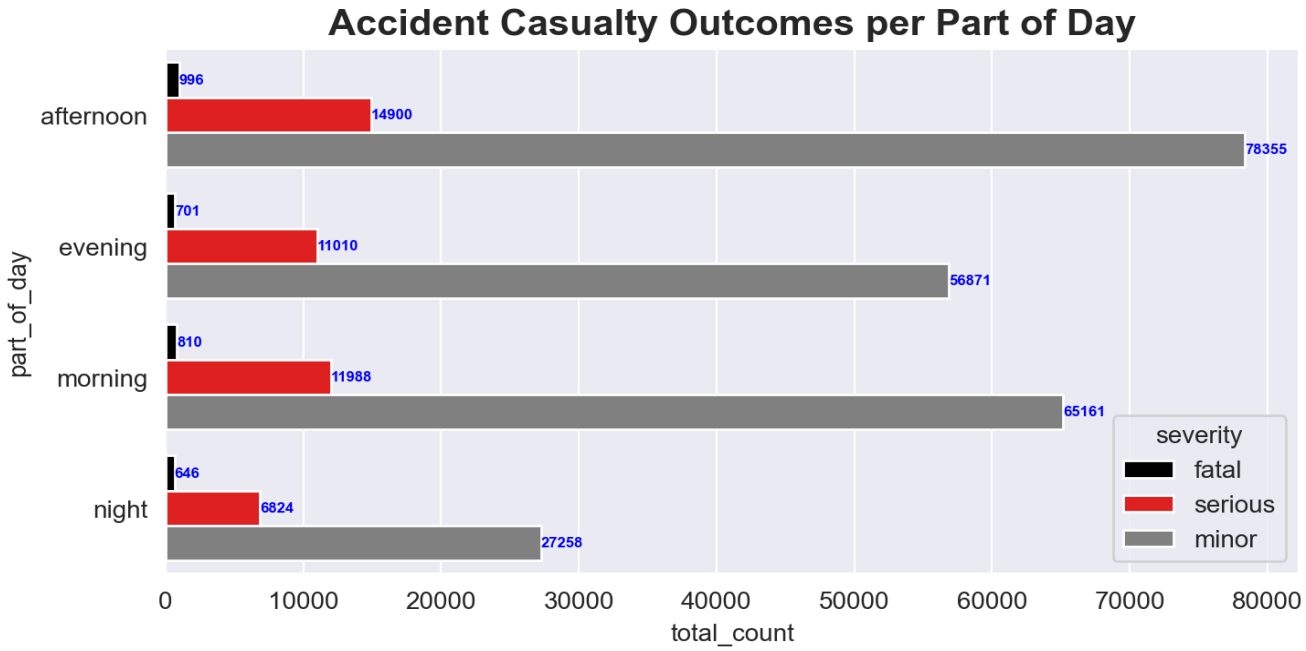

Association Between Variables
is_weekend
0
1
combined_variables
bus_or_coach_passenger: Not a bus or coach passenger & pedestrian_location: Not a Pedestrian & pedestrian_movement: Not a Pedestrian & pedestrian_crossing_human_control: None within 50 metres & junction_control: Give way or uncontrolled & carriageway_hazards: None & pedestrian_road_maintenance_worker: No / Not applicable & pedestrian_crossing_physical_facilities: No physical crossing facilities within 50 metres & 1st_road_class: Unclassified & road_type: Single carriageway
15.2%
41753
5.6%
15348
bus_or_coach_passenger: Not a bus or coach passenger & pedestrian_location: Not a Pedestrian & pedestrian_movement: Not a Pedestrian & pedestrian_crossing_human_control: None within 50 metres & junction_control: Give way or uncontrolled & carriageway_hazards: None & pedestrian_road_maintenance_worker: No / Not applicable & pedestrian_crossing_physical_facilities: No physical crossing facilities within 50 metres & 1st_road_class: B & road_type: Single carriageway
15878
bus_or_coach_passenger: Not a bus or coach passenger & pedestrian_location: Not a Pedestrian & pedestrian_movement: Not a Pedestrian & pedestrian_crossing_human_control: None within 50 metres & junction_control: Give way or uncontrolled & carriageway_hazards: None & pedestrian_road_maintenance_worker: No / Not applicable & pedestrian_crossing_physical_facilities: No physical crossing facilities within 50 metres & 1st_road_class: A & road_type: Dual carriageway
16227
bus_or_coach_passenger: Not a bus or coach passenger & pedestrian_location: Not a Pedestrian & pedestrian_movement: Not a Pedestrian & pedestrian_crossing_human_control: None within 50 metres & junction_control: Give way or uncontrolled & carriageway_hazards: None & pedestrian_road_maintenance_worker: No / Not applicable & pedestrian_crossing_physical_facilities: No physical crossing facilities within 50 metres & 1st_road_class: A & road_type: Single carriageway
14.5%
39933
0
20000
40000
60000
80000
100000
Accidents

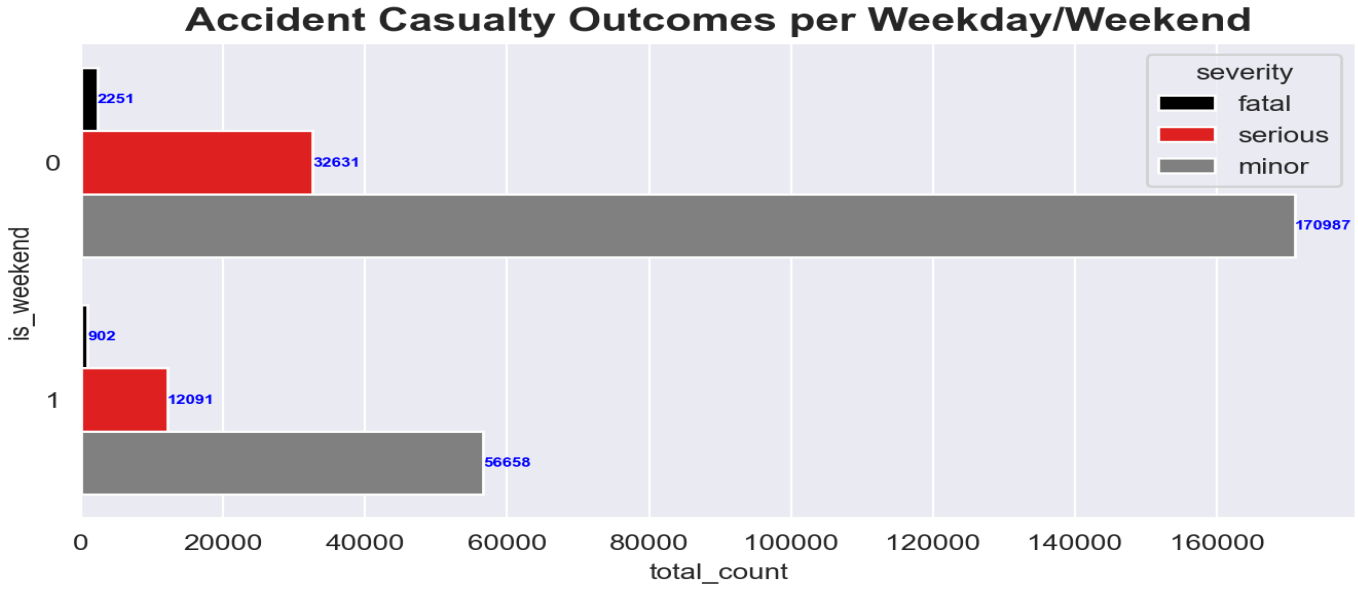

Accident Casualty Outcomes per road_surface_conditions
casualty_severity
minor
serious
fatal
Total_count
160000
140000
120000
100000
80000
60000
40000
20000
0
160069
58.1%
31225
11.3%
2204
299
56
4
2333
496
30
500
90
6
64444
23.4%
12855
4.7%
909
Dry
Flood over 3cm. deep
Frost or ice
Snow
Wet or damp
Road_surface_conditions


Accident Casualty Outcomes per road_type
casualty_severity
minor
serious
fatal
Road_type
Dual carriageway
46261
7206
666
One way street
3681
607
18
Roundabout
12783
1721
36
Single carriageway
57.8%
159236
12.6%
34579
2370
Slip road
4307
486
50
Unknown
1377
123
13
0
20000
40000
60000
80000
100000
120000
140000
160000
Total_count

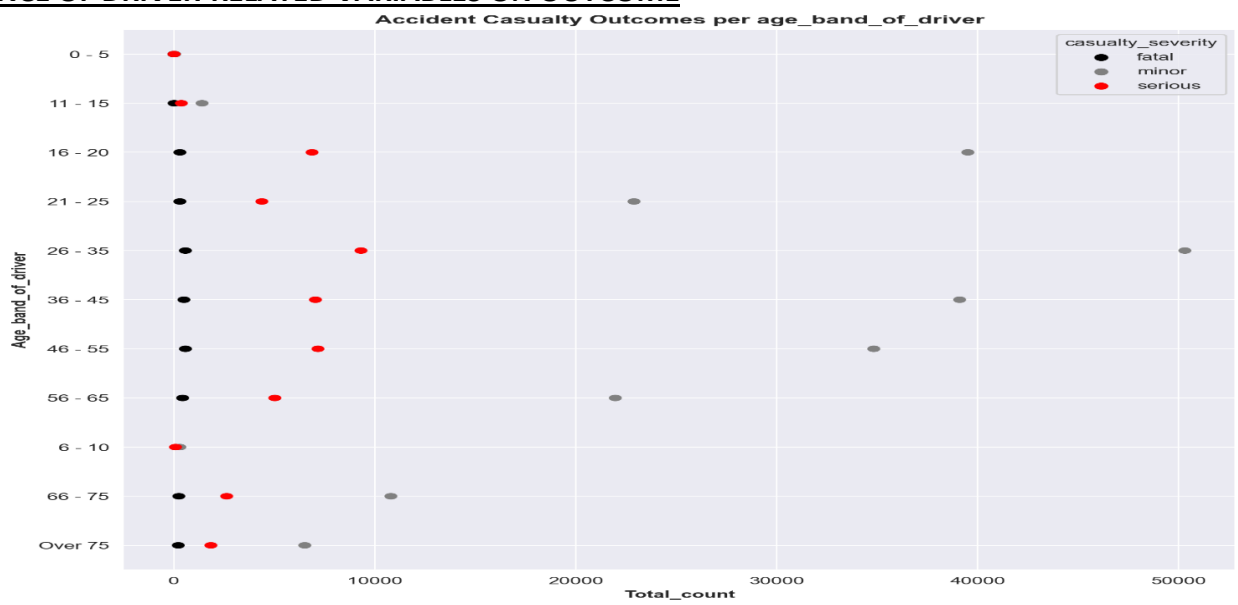

Accident Casualty Outcomes per special_conditions_at_site
Special_conditions_at_site
Auto signal part defective
20
80
Auto traffic signal - out
92
621
14
Mud
106
492
3
None
15.9%
43783
80.1%
222341
3103
Oil or diesel
100
361
7
Road sign or marking defective or obscured
57
442
6
Road surface defective
118
264
2
Roadworks
446
3044
18
casualty_severity
serious
minor
fatal
0
50000
100000
150000
200000
Total_count

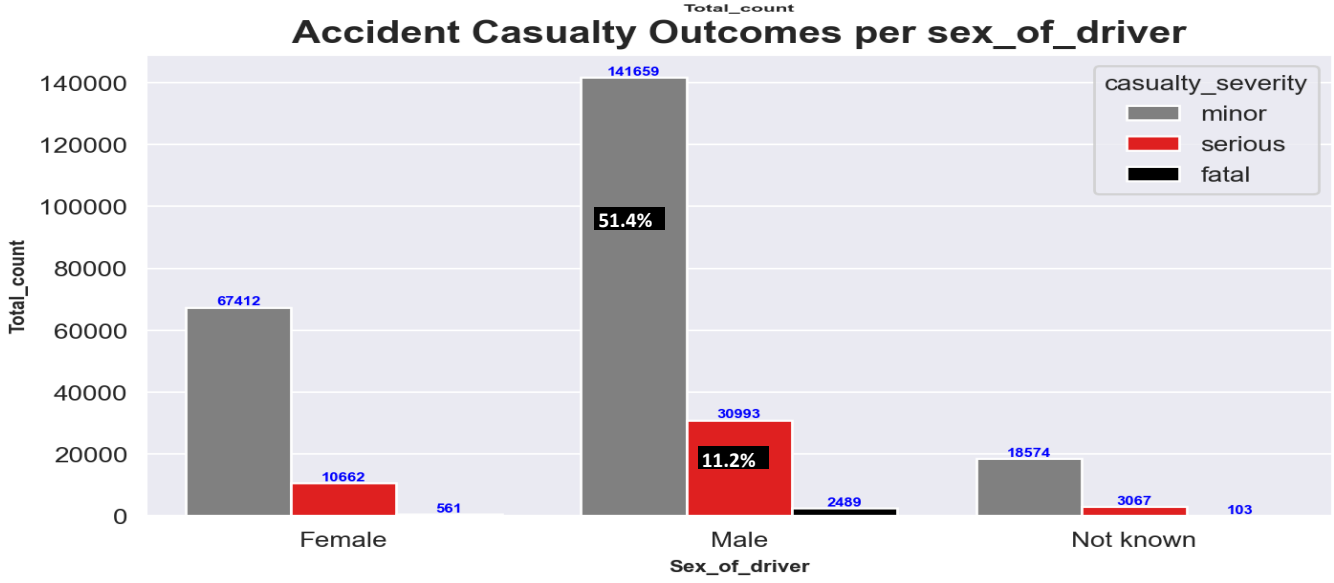

Accident Casualty Outcomes per speed_limit
casualty_severity
fatal
serious
minor
Speed_limit
20 MPH
30 MPH
40 MPH
50 MPH
60 MPH
70 MPH
75
2839
16160
885
22559
123788
341
4359
22663
235
2302
11901
1146
9168
30919
471
3495
22214
0
20000
40000
60000
80000
100000
120000
Total_count

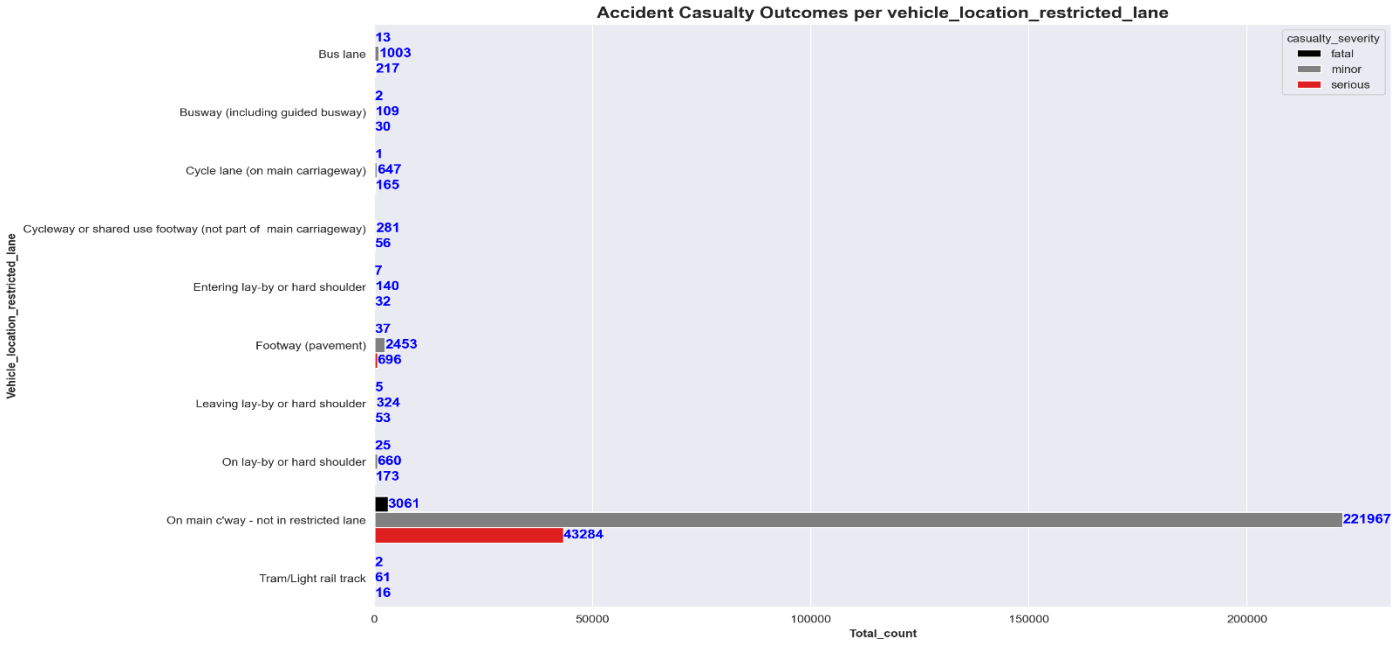

Accident Casualty Outcomes per Part of Day
part_of_day
afternoon
evening
morning
night
996
14900
78355
701
11010
56871
810
11988
65161
646
6824
27258
severity
fatal
serious
minor
0
10000
20000
30000
40000
50000
60000
70000
80000
total_count

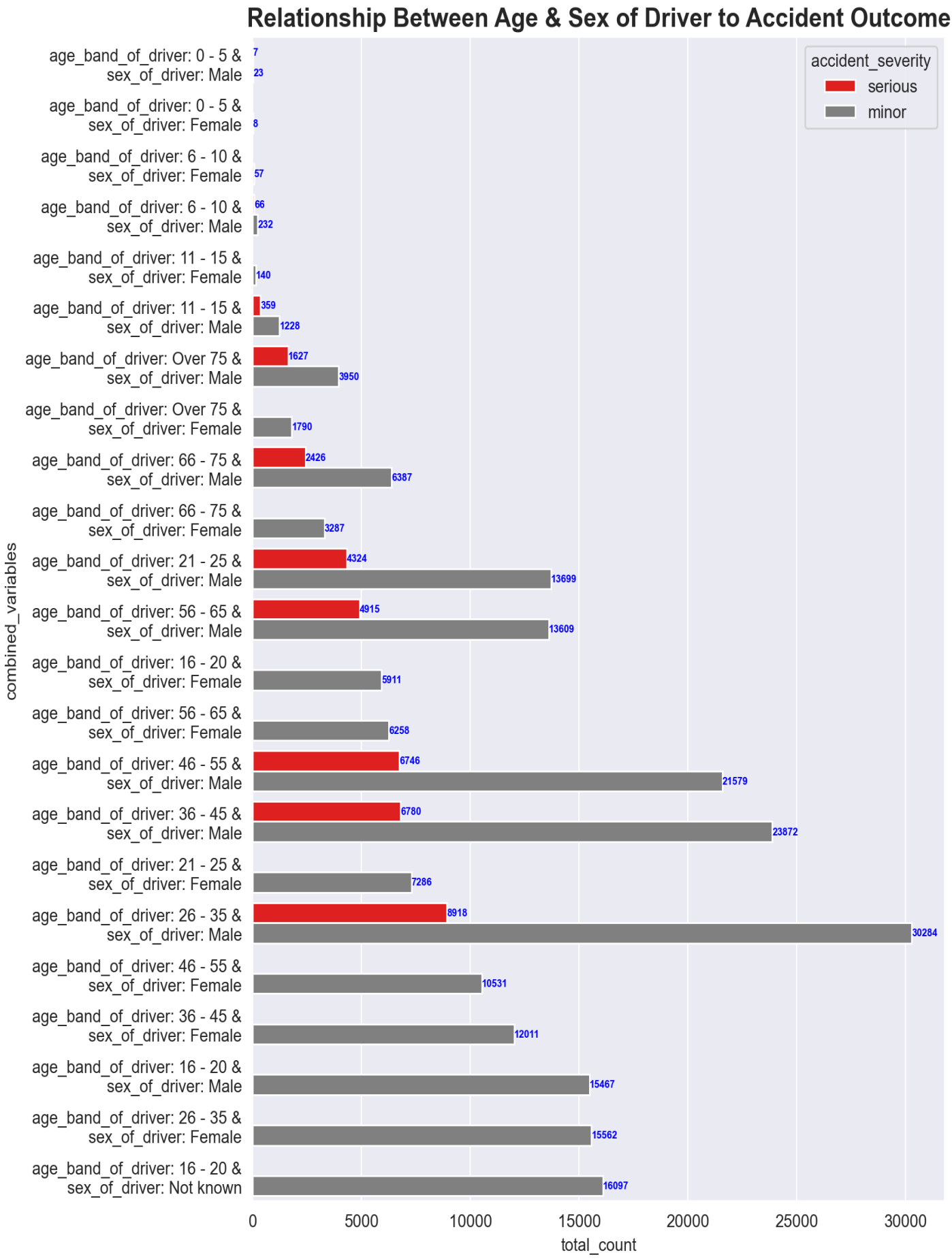

Accident Casualty Outcomes per Weekday/Weekend
is_weekend
0
1
2251
32631
170987
902
12091
56658
severity
fatal
serious
minor
0
20000
40000
60000
80000
100000
120000
140000
160000
total_count

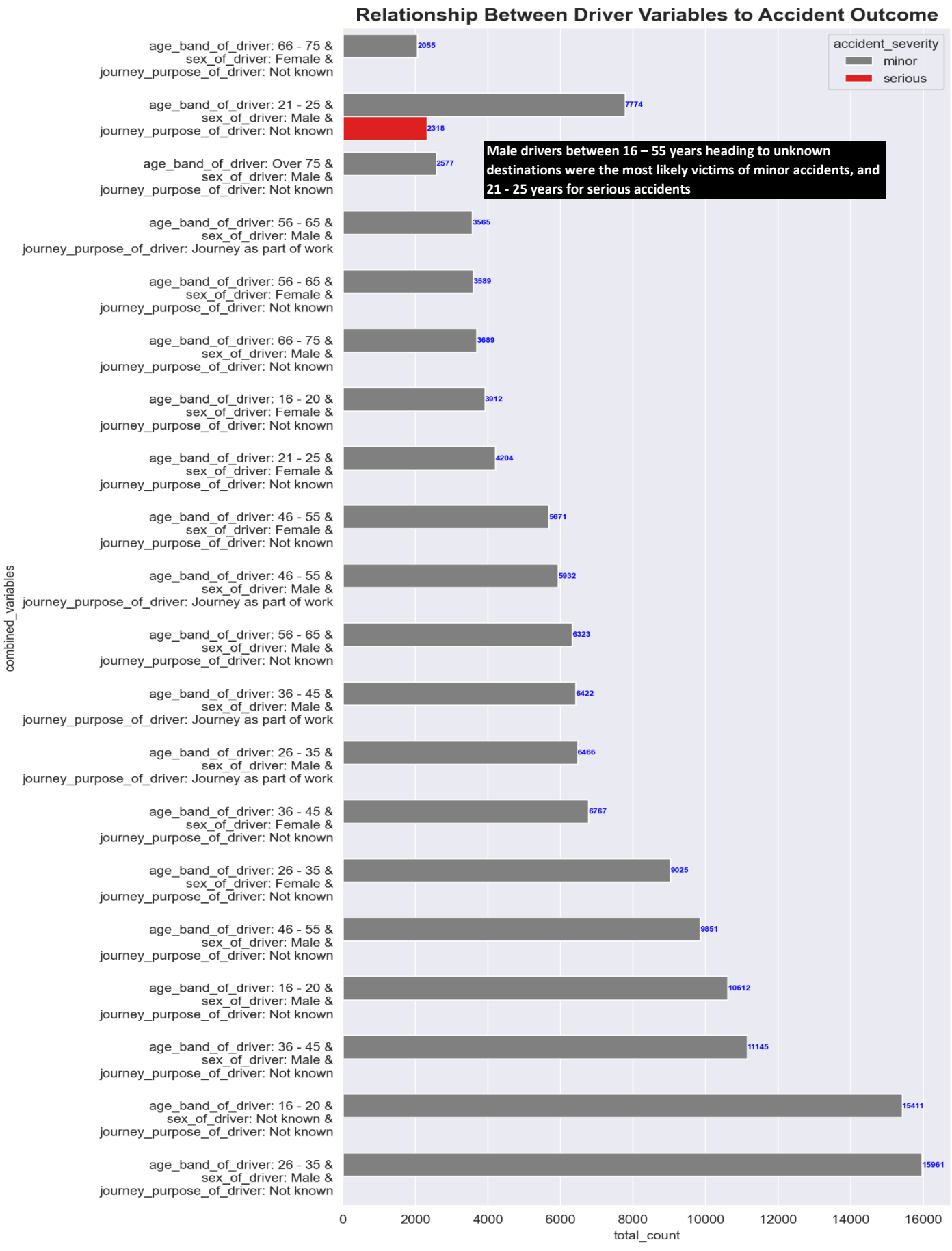
INFLUENCE OF DRIVER-RELATED VARIABLES ON OUTCOME

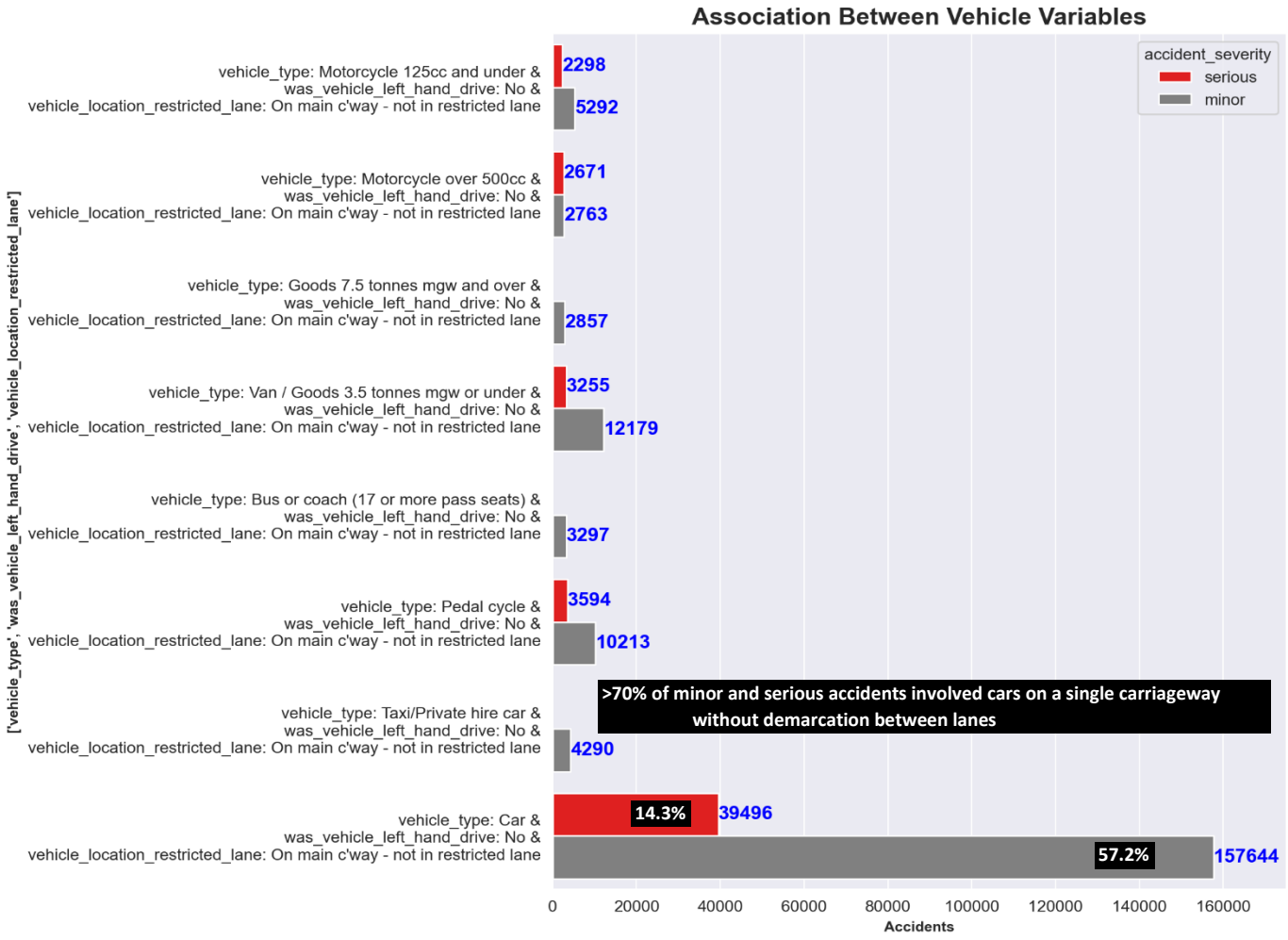


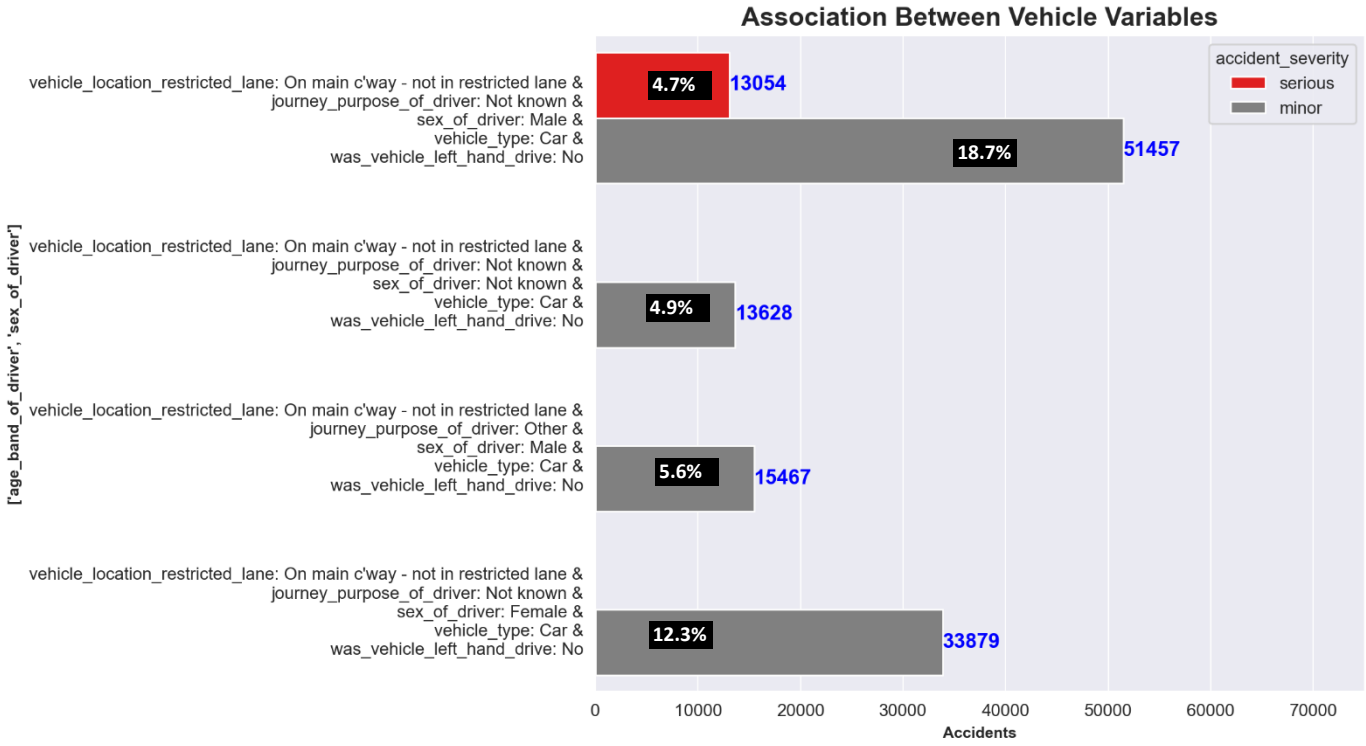




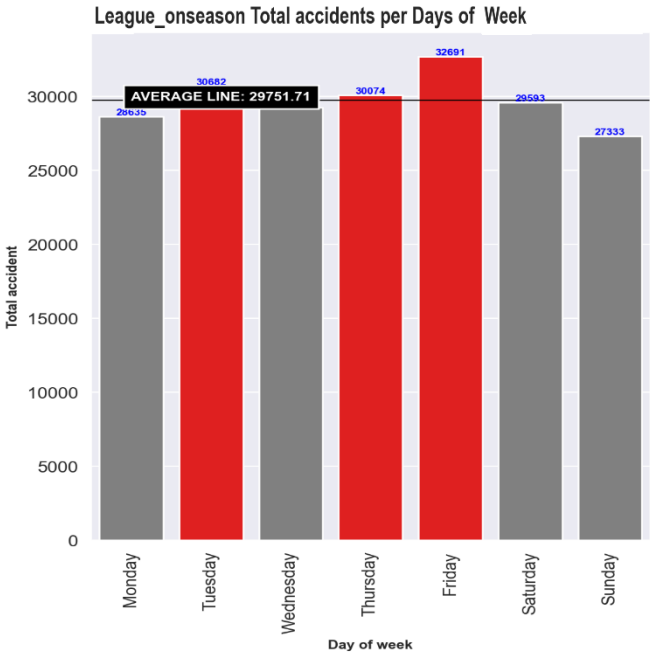

Relationship Between Age & Sex of Driver to Accident Outcome
accident_severity
serious
minor
combined_variables
age_band_of_driver: 0 - 5 & sex_of_driver: Male
age_band_of_driver: 0 - 5 & sex_of_driver: Female
age_band_of_driver: 6 - 10 & sex_of_driver: Female
age_band_of_driver: 6 - 10 & sex_of_driver: Male
age_band_of_driver: 11 - 15 & sex_of_driver: Female
age_band_of_driver: 11 - 15 & sex_of_driver: Male
age_band_of_driver: Over 75 & sex_of_driver: Male
age_band_of_driver: Over 75 & sex_of_driver: Female
age_band_of_driver: 66 - 75 & sex_of_driver: Male
age_band_of_driver: 66 - 75 & sex_of_driver: Female
age_band_of_driver: 21 - 25 & sex_of_driver: Male
age_band_of_driver: 56 - 65 & sex_of_driver: Male
age_band_of_driver: 16 - 20 & sex_of_driver: Female
age_band_of_driver: 56 - 65 & sex_of_driver: Female
age_band_of_driver: 46 - 55 & sex_of_driver: Male
age_band_of_driver: 36 - 45 & sex_of_driver: Male
age_band_of_driver: 21 - 25 & sex_of_driver: Female
age_band_of_driver: 26 - 35 & sex_of_driver: Male
age_band_of_driver: 46 - 55 & sex_of_driver: Female
age_band_of_driver: 36 - 45 & sex_of_driver: Female
age_band_of_driver: 16 - 20 & sex_of_driver: Male
age_band_of_driver: 26 - 35 & sex_of_driver: Female
age_band_of_driver: 16 - 20 & sex_of_driver: Not known
7
23
8
57
66
232
140
359
1228
1627
3950
1790
2426
6387
3287
4324
13699
4915
13609
5911
6258
6746
21579
6780
23872
7286
8918
30284
10531
12011
15467
15562
16097
0
5000
10000
15000
20000
25000
30000
total_count

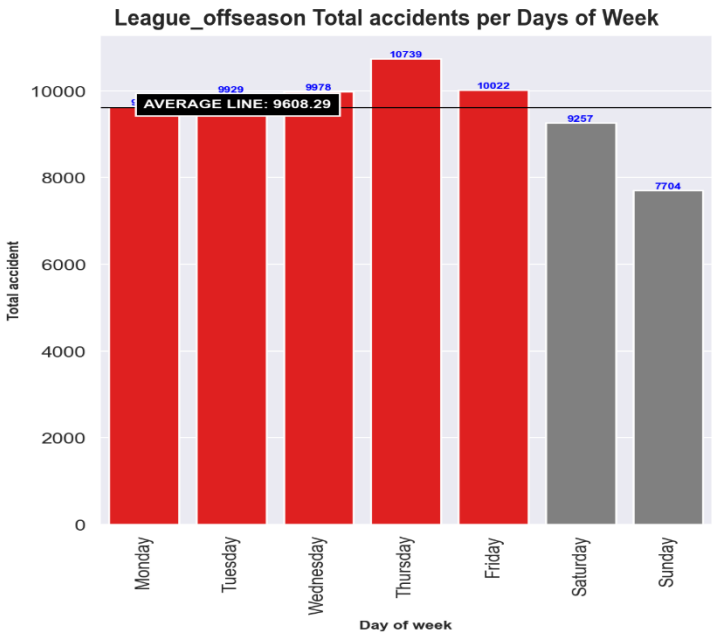

Relationship Between Driver Variables to Accident Outcome
accident_severity
minor
serious
Male drivers between 16 – 55 years heading to unknown destinations were the most likely victims of minor accidents, and 21 - 25 years for serious accidents
age_band_of_driver: 66 - 75 & sex_of_driver: Female & journey_purpose_of_driver: Not known
2055
age_band_of_driver: 21 - 25 & sex_of_driver: Male & journey_purpose_of_driver: Not known
7774
2318
age_band_of_driver: Over 75 & sex_of_driver: Male & journey_purpose_of_driver: Not known
2577
age_band_of_driver: 56 - 65 & sex_of_driver: Male & journey_purpose_of_driver: Journey as part of work
3565
age_band_of_driver: 56 - 65 & sex_of_driver: Female & journey_purpose_of_driver: Not known
3589
age_band_of_driver: 66 - 75 & sex_of_driver: Male & journey_purpose_of_driver: Not known
3689
age_band_of_driver: 16 - 20 & sex_of_driver: Female & journey_purpose_of_driver: Not known
3912
age_band_of_driver: 21 - 25 & sex_of_driver: Female & journey_purpose_of_driver: Not known
4204
age_band_of_driver: 46 - 55 & sex_of_driver: Female & journey_purpose_of_driver: Not known
5671
age_band_of_driver: 46 - 55 & sex_of_driver: Male & journey_purpose_of_driver: Journey as part of work
5932
age_band_of_driver: 56 - 65 & sex_of_driver: Male & journey_purpose_of_driver: Not known
6323
age_band_of_driver: 36 - 45 & sex_of_driver: Male & journey_purpose_of_driver: Journey as part of work
6422
age_band_of_driver: 26 - 35 & sex_of_driver: Male & journey_purpose_of_driver: Journey as part of work
6466
age_band_of_driver: 36 - 45 & sex_of_driver: Female & journey_purpose_of_driver: Not known
6767
age_band_of_driver: 26 - 35 & sex_of_driver: Female & journey_purpose_of_driver: Not known
9025
age_band_of_driver: 46 - 55 & sex_of_driver: Male & journey_purpose_of_driver: Not known
9851
age_band_of_driver: 16 - 20 & sex_of_driver: Male & journey_purpose_of_driver: Not known
10612
age_band_of_driver: 36 - 45 & sex_of_driver: Male & journey_purpose_of_driver: Not known
11145
age_band_of_driver: 16 - 20 & sex_of_driver: Not known & journey_purpose_of_driver: Not known
15411
age_band_of_driver: 26 - 35 & sex_of_driver: Male & journey_purpose_of_driver: Not known
15961
combined_variables
0
2000
4000
6000
8000
10000
12000
14000
16000
total_count

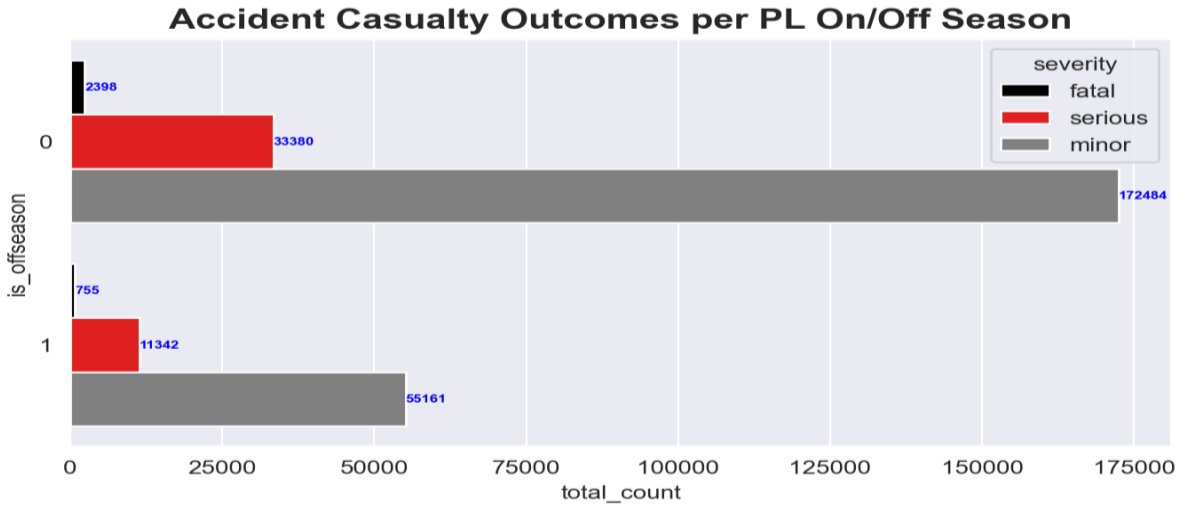

Association Between Vehicle Variables
accident_severity
serious
minor
['vehicle_type', 'was_vehicle_left_hand_drive', 'vehicle_location_restricted_lane']
vehicle_type: Motorcycle 125cc and under & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
2298
5292
vehicle_type: Motorcycle over 500cc & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
2671
2763
vehicle_type: Goods 7.5 tonnes mgw and over & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
2857
vehicle_type: Van / Goods 3.5 tonnes mgw or under & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
3255
12179
vehicle_type: Bus or coach (17 or more pass seats) & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
3297
vehicle_type: Pedal cycle & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
3594
10213
>70% of minor and serious accidents involved cars on a single carriageway without demarcation between lanes
vehicle_type: Taxi/Private hire car & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
4290
vehicle_type: Car & was_vehicle_left_hand_drive: No & vehicle_location_restricted_lane: On main c'way - not in restricted lane
14.3%
39496
57.2%
157644
0
20000
40000
60000
80000
100000
120000
140000
160000
Accidents

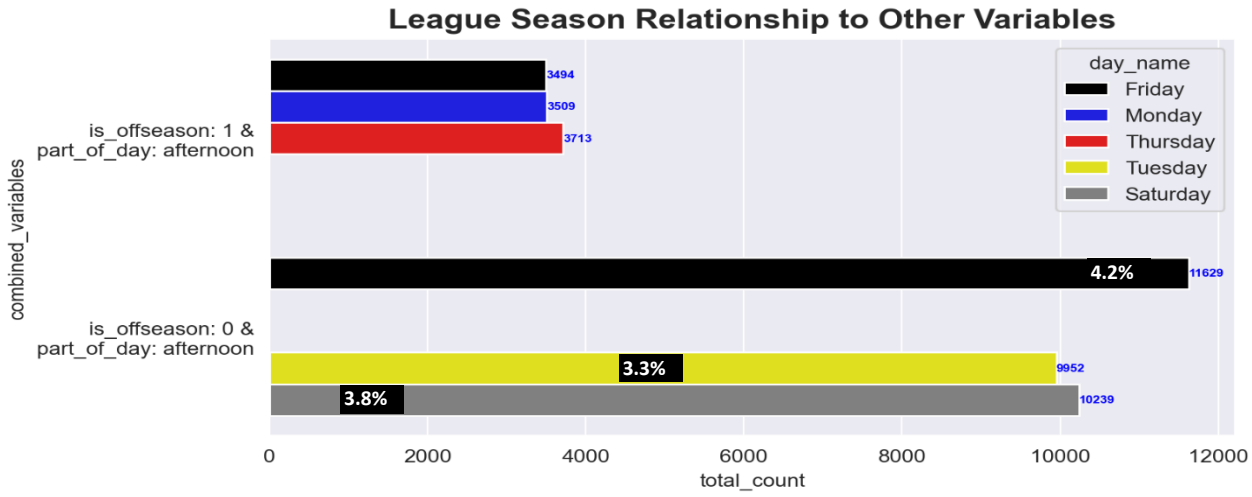

Association Between Vehicle Variables
accident_severity
serious
minor
['age_band_of_driver', 'sex_of_driver']
vehicle_location_restricted_lane: On main c'way - not in restricted lane & journey_purpose_of_driver: Not known & sex_of_driver: Male & vehicle_type: Car & was_vehicle_left_hand_drive: No
4.7%
13054
18.7%
51457
vehicle_location_restricted_lane: On main c'way - not in restricted lane & journey_purpose_of_driver: Not known & sex_of_driver: Not known & vehicle_type: Car & was_vehicle_left_hand_drive: No
4.9%
13628
vehicle_location_restricted_lane: On main c'way - not in restricted lane & journey_purpose_of_driver: Other & sex_of_driver: Male & vehicle_type: Car & was_vehicle_left_hand_drive: No
5.6%
15467
vehicle_location_restricted_lane: On main c'way - not in restricted lane & journey_purpose_of_driver: Not known & sex_of_driver: Female & vehicle_type: Car & was_vehicle_left_hand_drive: No
12.3%
33879
0
10000
20000
30000
40000
50000
60000
70000
Accidents

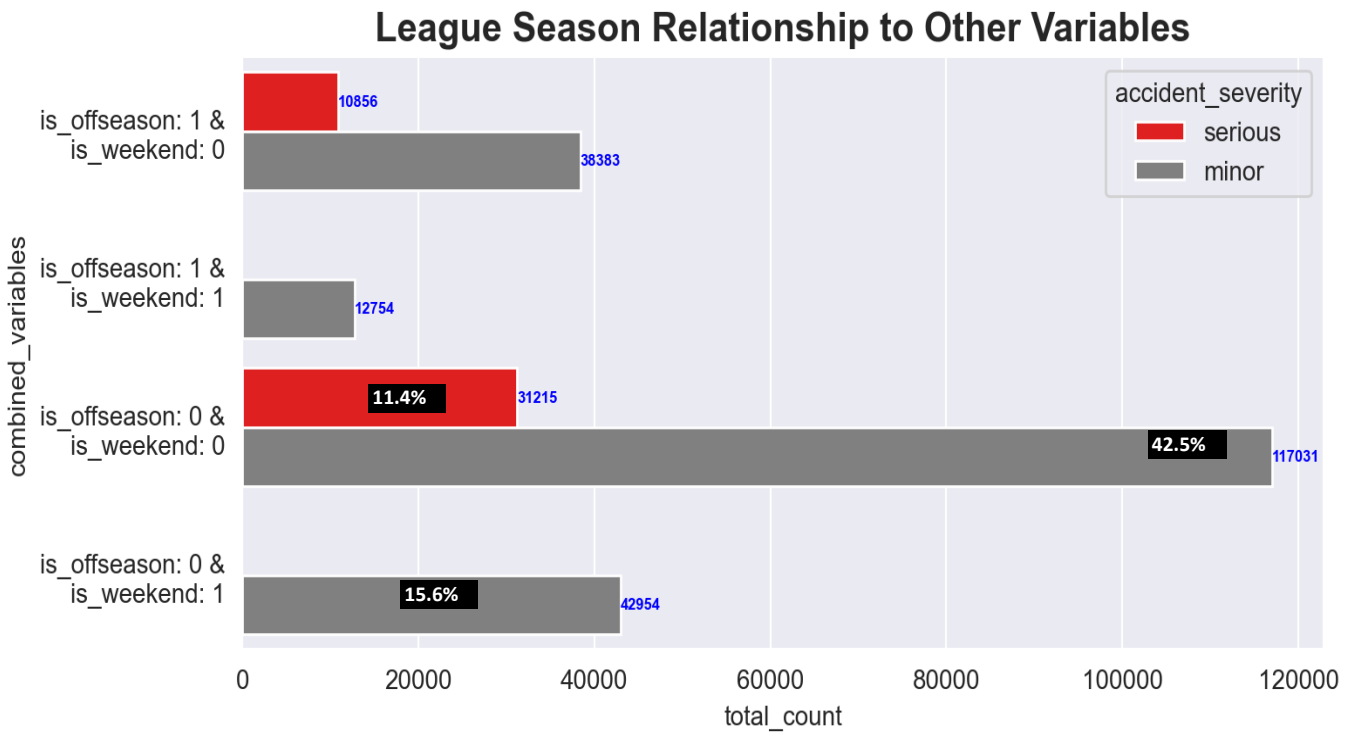
PREMIER LEAGUE SEASON ACCIDENTS

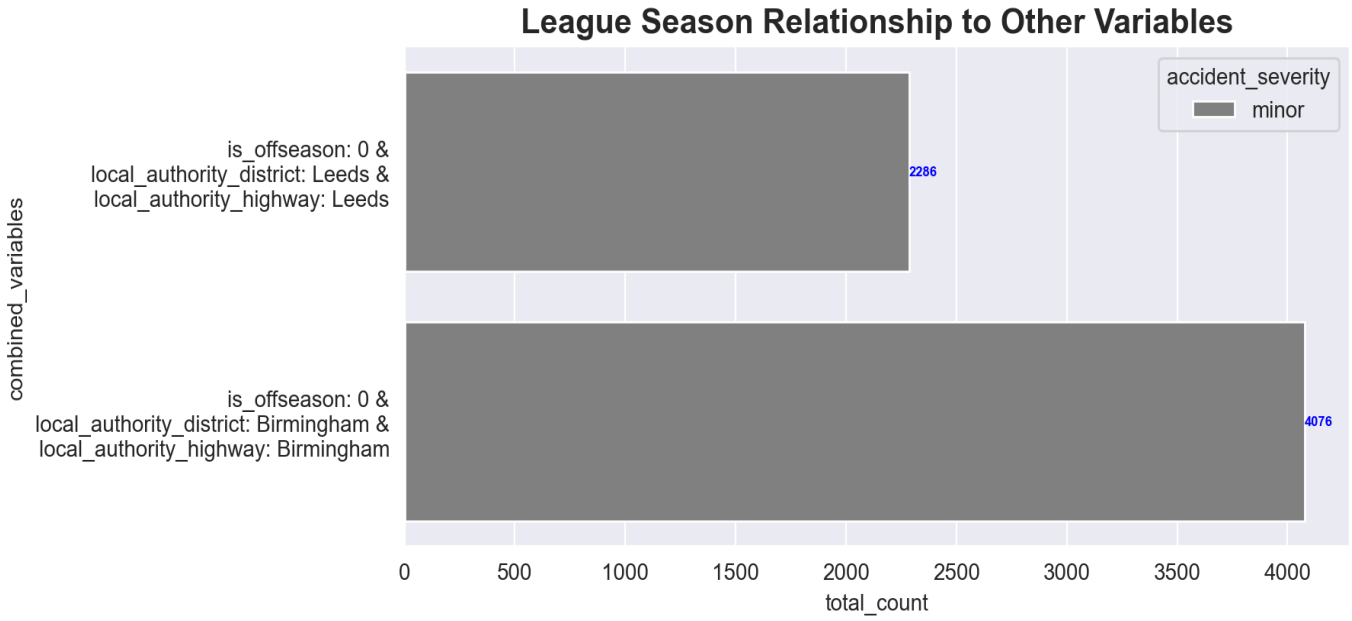


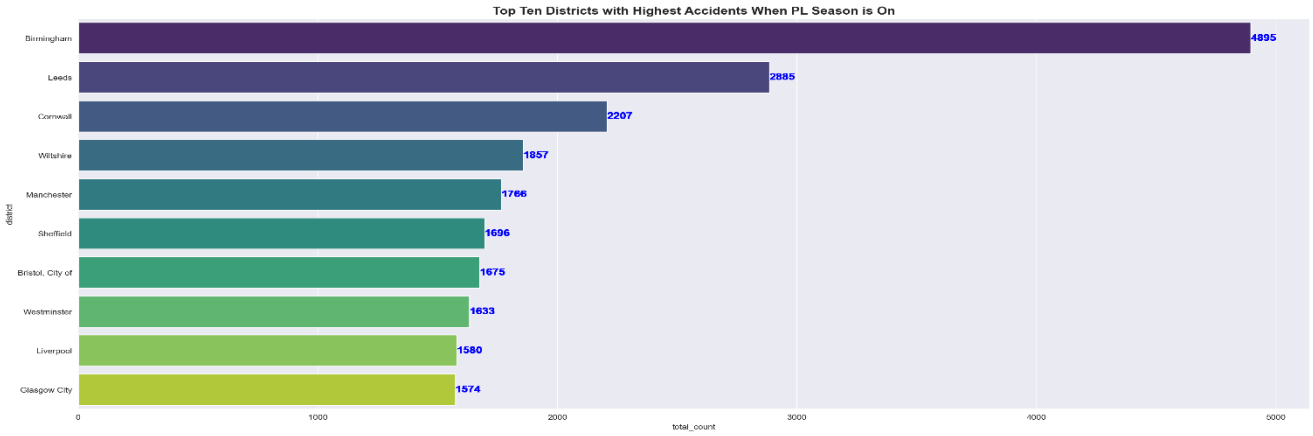


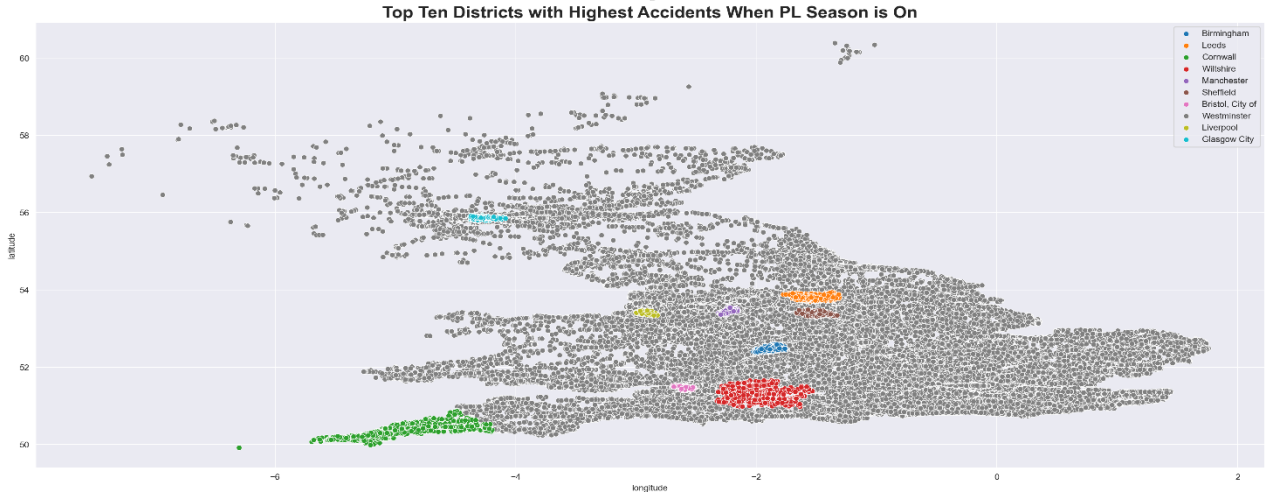


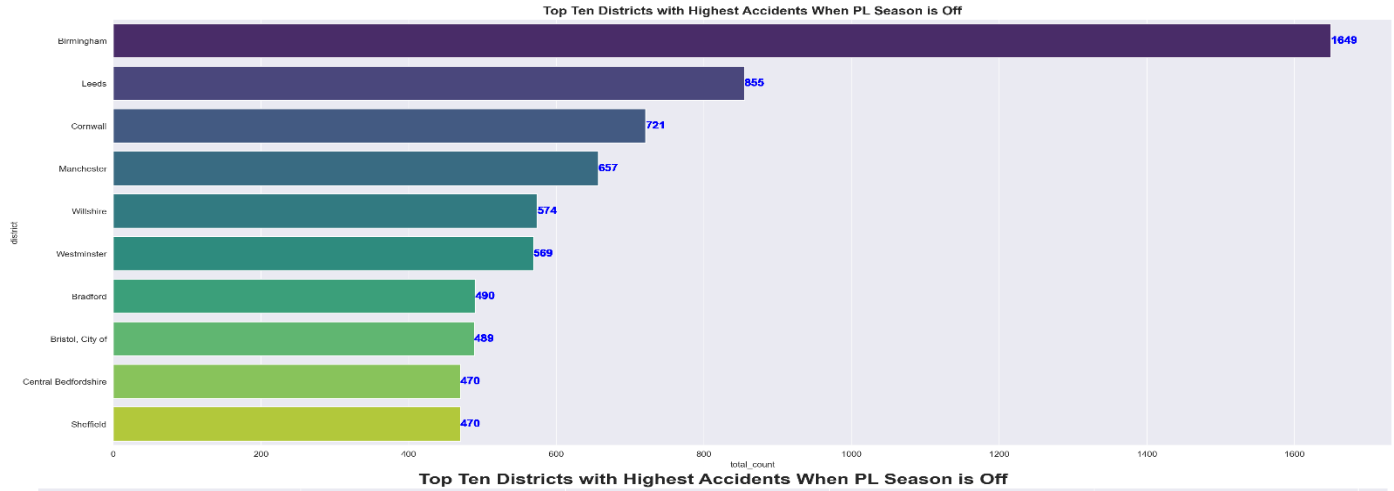


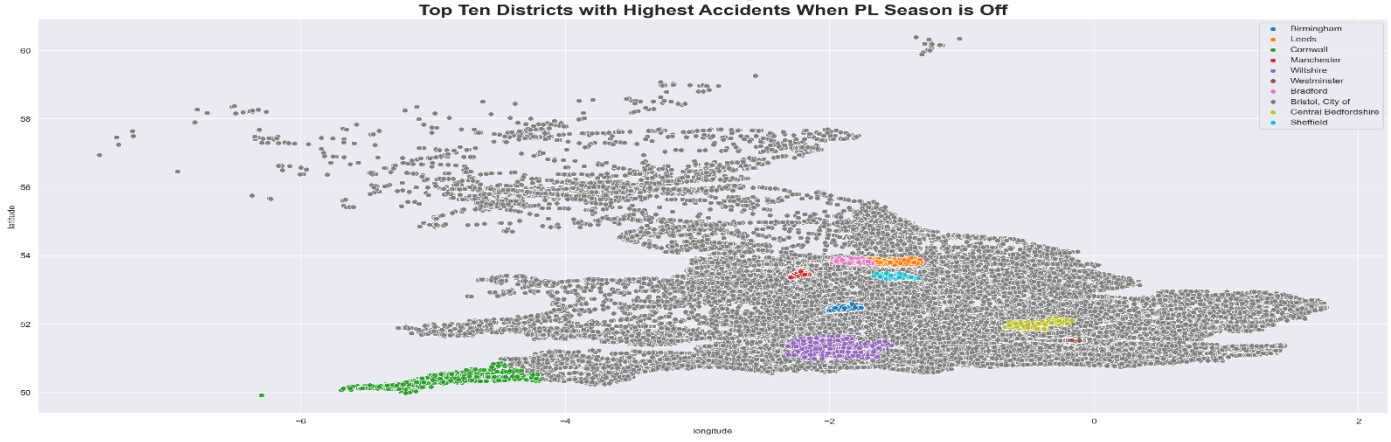

League Season Relationship to Other Variables
accident_severity
serious
minor
combined_variables
is_offseason: 1 & is_weekend: 0
is_offseason: 1 & is_weekend: 1
is_offseason: 0 & is_weekend: 0
is_offseason: 0 & is_weekend: 1
10856
38383
12754
11.4%
31215
42.5%
117031
15.6%
42954
0
20000
40000
60000
80000
100000
120000
total_count


League Season Relationship to Other Variables
accident_severity
minor
combined_variables
is_offseason: 0 & local_authority_district: Leeds & local_authority_highway: Leeds
is_offseason: 0 & local_authority_district: Birmingham & local_authority_highway: Birmingham
2286
4076
0
500
1000
1500
2000
2500
3000
3500
4000
total_count


Top Ten Districts with Highest Accidents When PL Season is On
Birmingham
Leeds
Cornwall
Wiltshire
Manchester
Sheffield
Bristol, City of
Westminster
Liverpool
Glasgow City
4895
2885
2207
1857
1766
1696
1675
1633
1580
1574
district
total_count
0
1000
2000
3000
4000
5000


Top Ten Districts with Highest Accidents When PL Season is On
Birmingham
Leeds
Cornwall
Wiltshire
Manchester
Sheffield
Bristol, City of
Westminster
Liverpool
Glasgow City
latitude
longitude
60
58
56
54
52
50
−6
−4
−2
0
2


Top Ten Districts with Highest Accidents When PL Season is Off
Birmingham
Leeds
Cornwall
Manchester
Willshire
Westminster
Bradford
Bristol, City of
Central Bedfordshire
Sheffield
1649
855
721
657
574
569
490
489
470
470
district
total_count
0
200
400
600
800
1000
1200
1400
1600

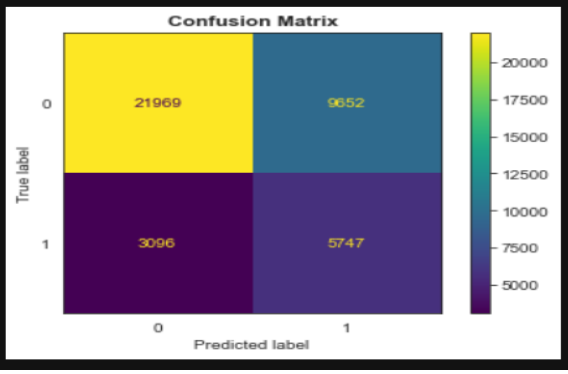

Top Ten Districts with Highest Accidents When PL Season is Off
Birmingham
Leeds
Cornwall
Manchester
Willshire
Westminster
Bradford
Bristol, City of
Central Bedfordshire
Sheffield
latitude
longitude
60
58
56
54
52
50
−6
−4
−2
0
2

PREDICTIONS

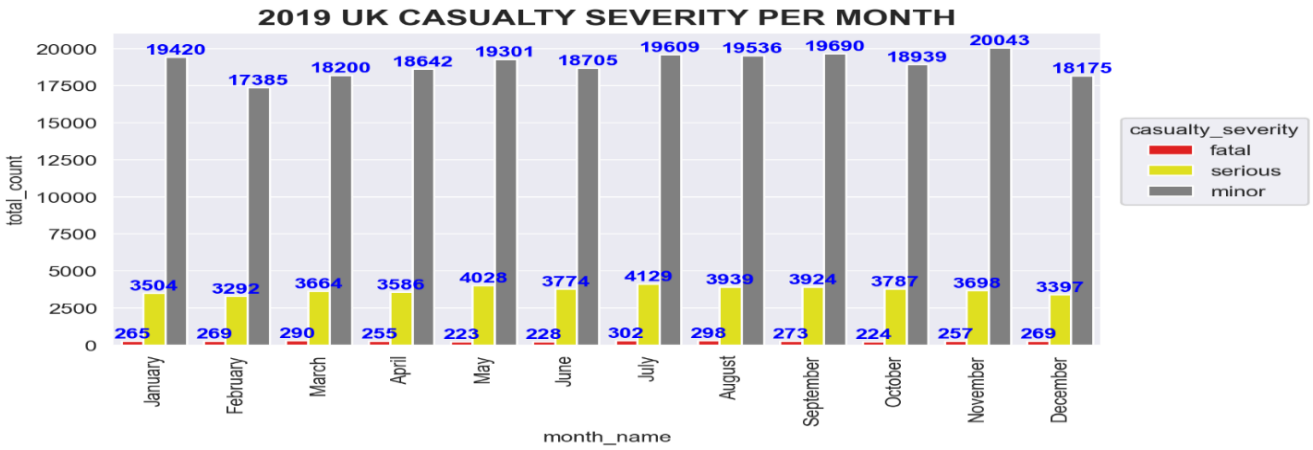
PREDICTING ACCIDENT SEVERITY

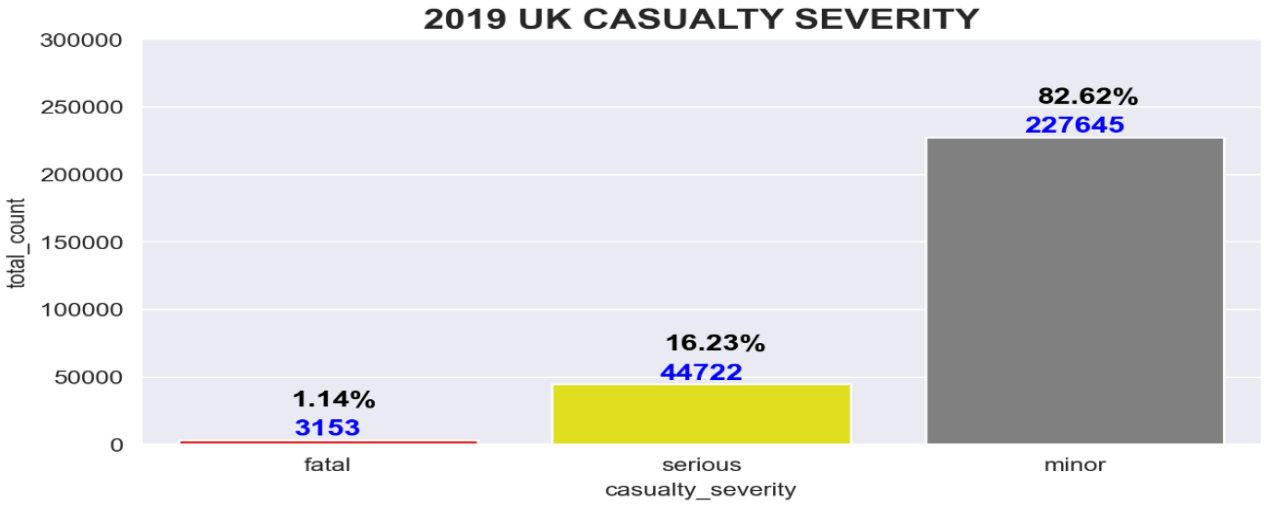
	precision	recall	f1-score	support
0	0.88	0.69	0.78	31621
1	0.37	0.65	0.47	8843
accuracy			0.68	40464
macro avg	0.62	0.67	0.62	40464
weighted avg	0.77	0.68	0.71	40464

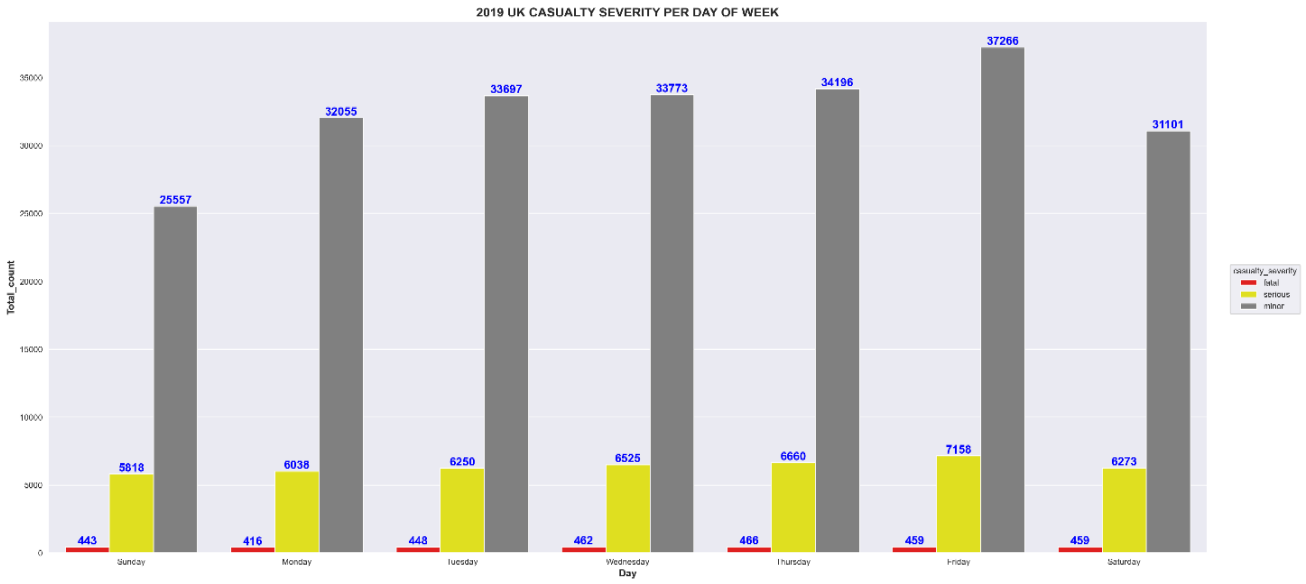
```
35]: <sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x14b32f7ac10>
```

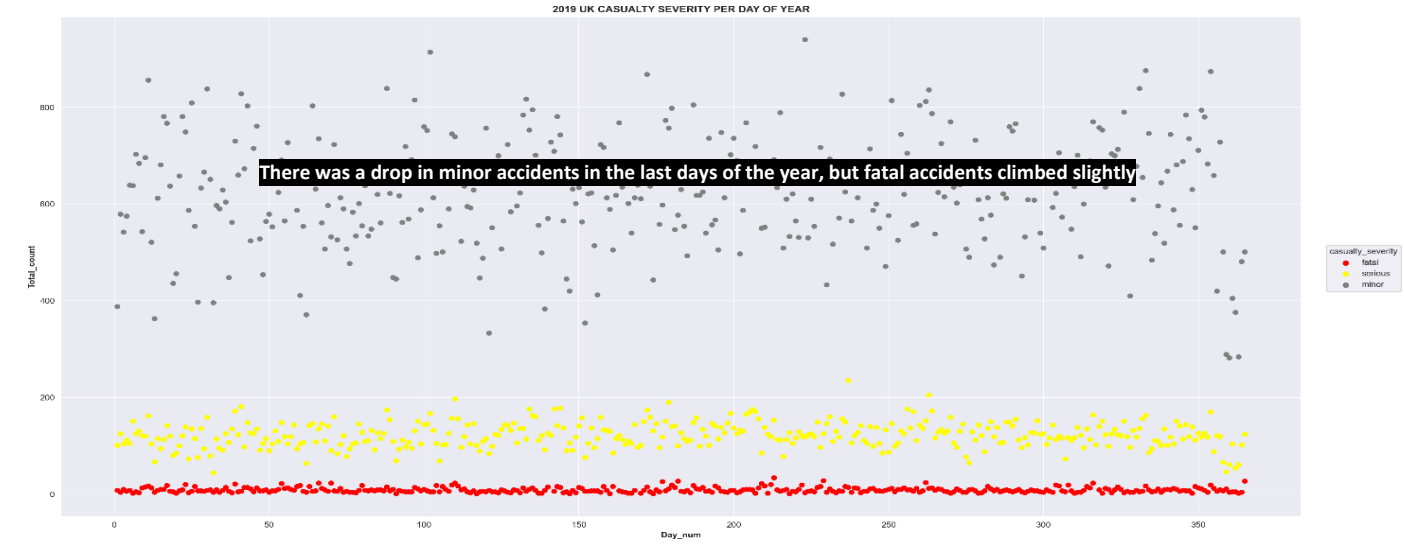


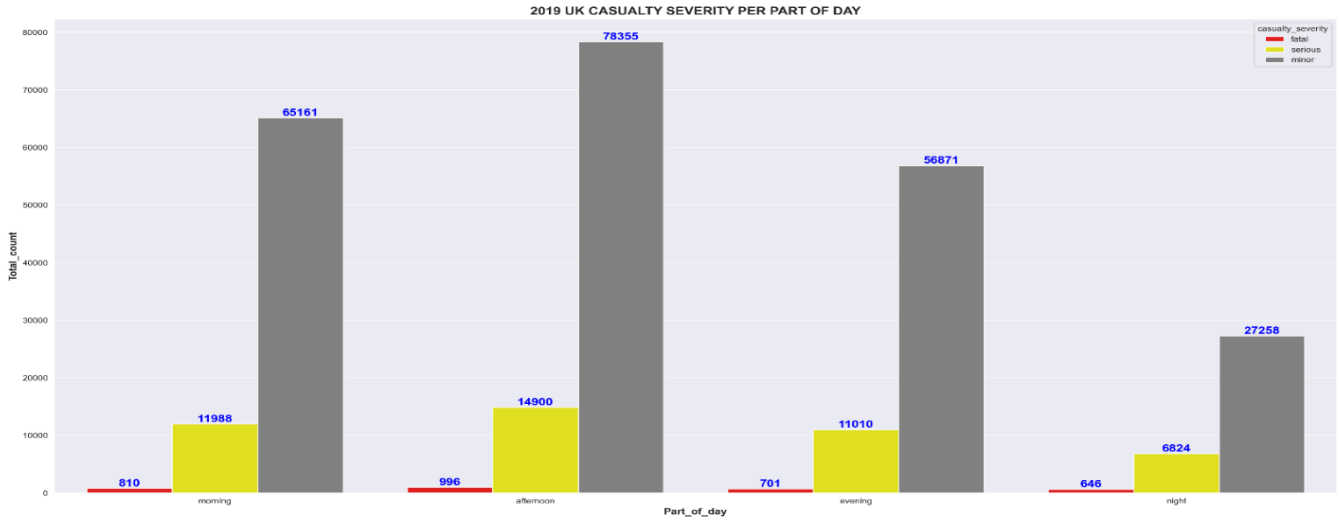
CASUALTY SEVERITY STATS

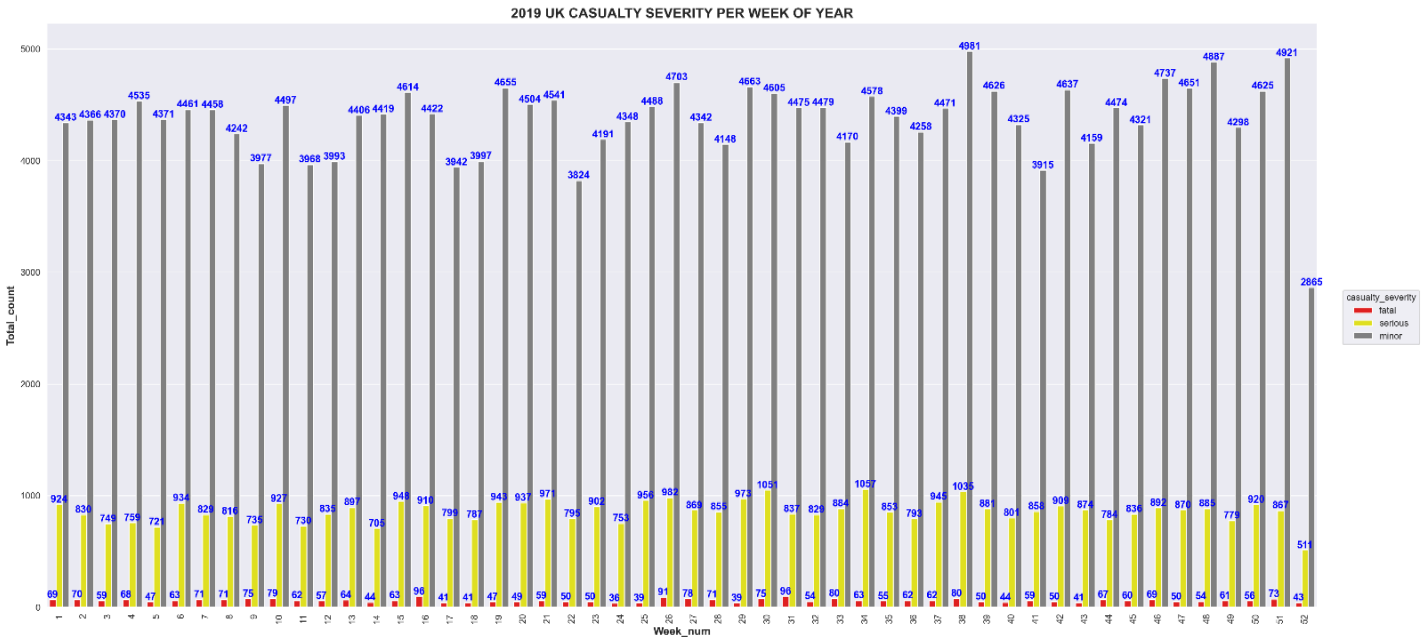





2019 UK CASUALTY SEVERITY PER DAY OF WEEK
Total_count
Day
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
443
5818
25557
416
6038
32055
448
6250
33697
462
6525
33773
466
6660
34196
459
7158
37266
459
6273
31101
casualty_severity
fatal
serious
minor


2019 UK CASUALTY SEVERITY PER DAY OF YEAR
There was a drop in minor accidents in the last days of the year, but fatal accidents climbed slightly
Total_count
Day_num
casualty_severity
fatal
serious
minor


2019 UK CASUALTY SEVERITY PER PART OF DAY
Total_count
Part_of_day
morning
afternoon
evening
night
810
11988
65161
996
14900
78355
701
11010
56871
646
6824
27258
casualty_severity
fatal
serious
minor



RECOMMENDATIONS

1. More campaigns should be taken to venues like football stadiums and sports centers in the top ten districts to educate young males about dangers of over-speeding, drunk driving, and driving at sunrise and sunset hours.
2. Always enforce a strict adherence to the speed limit, penalize defaulters and test the alcohol/drug level of suspected drivers before allowing them back on the road especially on Friday, Saturday and Sunday nights.
3. Increase the speed bumps and add minor road demarcations (where possible) to the busiest non-restricted single carriageways in places like Birmingham, Kent, and Cornwall.
4. Ensure that all streetlights along public roads are in good conditions to aid night visibility and increase traffic controls (human/facility) on single carriageways.