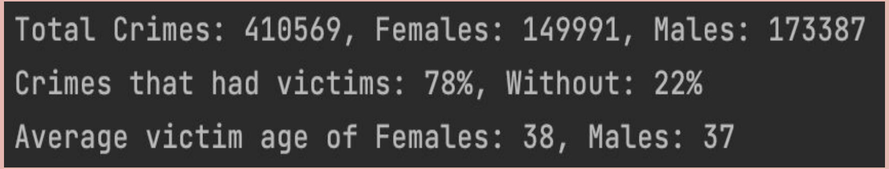
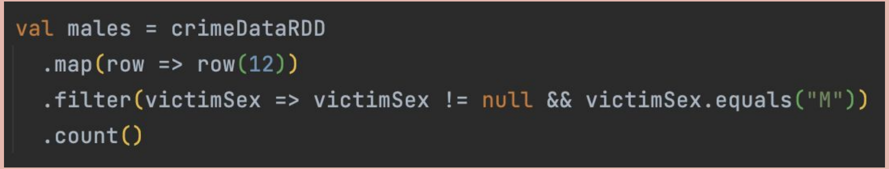
LA County Crime Trends

By Frank Tucci, Luke Franks, Winton Gee, Victor Phan

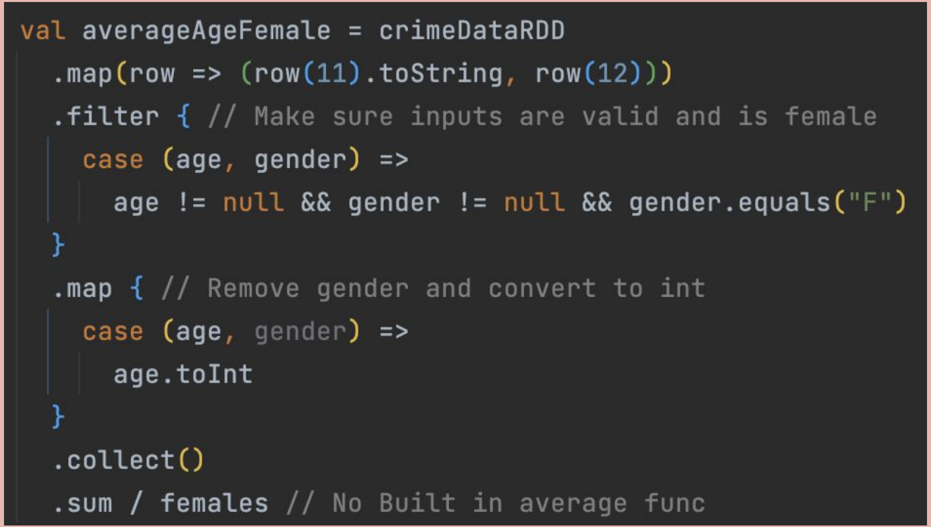
Victim Demographics



```
Total Crimes: 410569, Females: 149991, Males: 173387
Crimes that had victims: 78%, Without: 22%
Average victim age of Females: 38, Males: 37
```

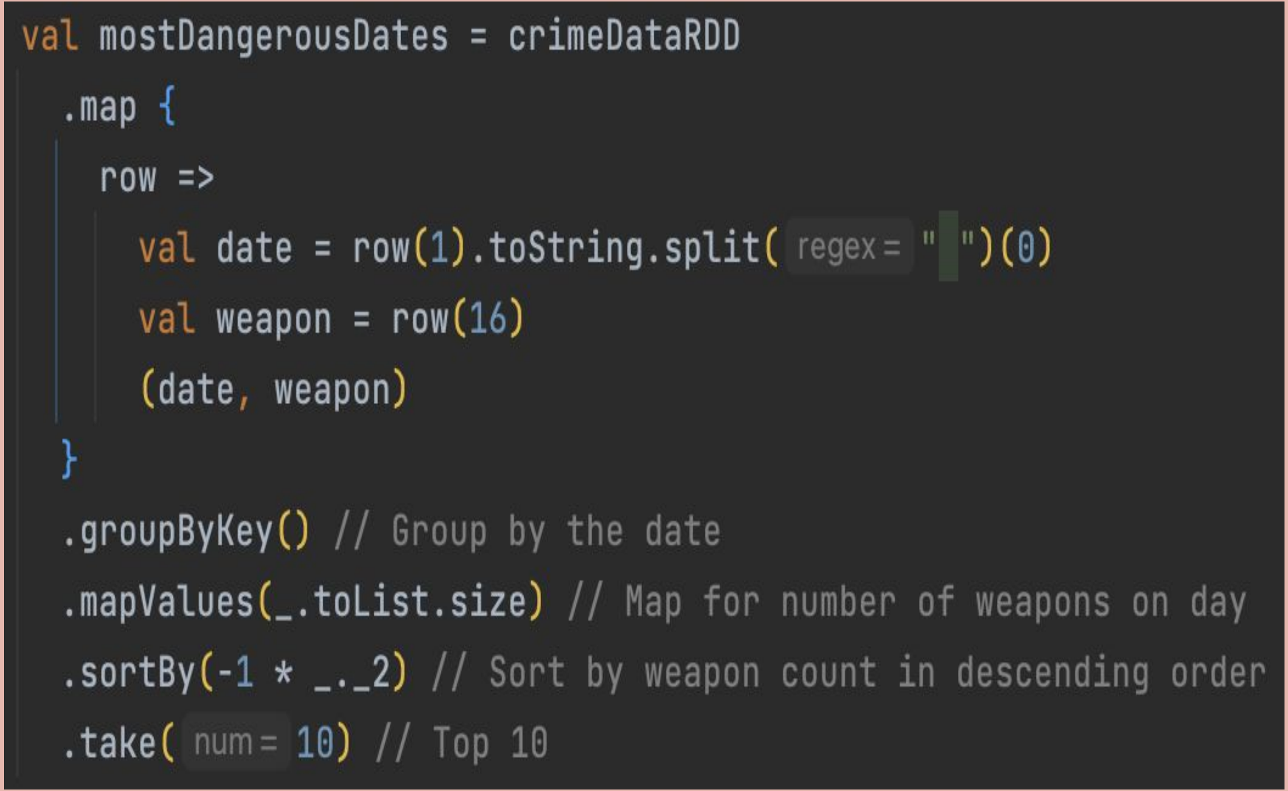


```scala
val males = crimeDataRDD
  .map(row => row(12))
  .filter(victimSex => victimSex != null && victimSex.equals("M"))
  .count()
```

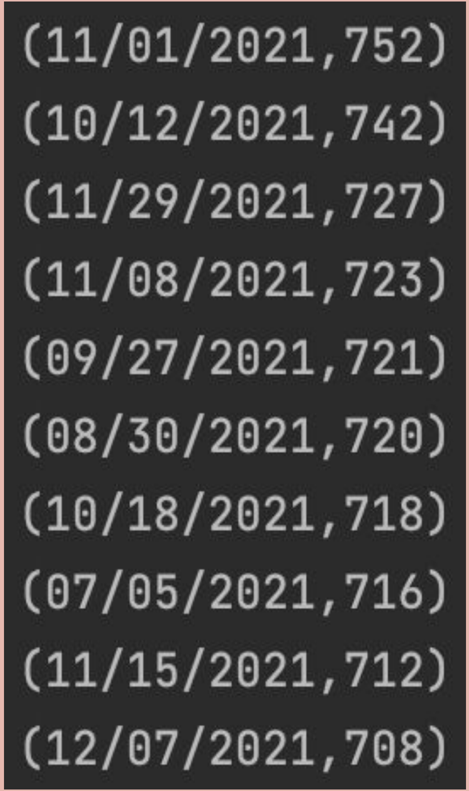


```scala
val averageAgeFemale = crimeDataRDD
  .map(row => (row(11).toString, row(12)))
  .filter { // Make sure inputs are valid and is female
    case (age, gender) =>
      age != null && gender != null && gender.equals("F")
  }
  .map { // Remove gender and convert to int
    case (age, gender) =>
      age.toInt
  }
  .collect()
  .sum / females // No Built in average func
```

Most Dangerous Dates

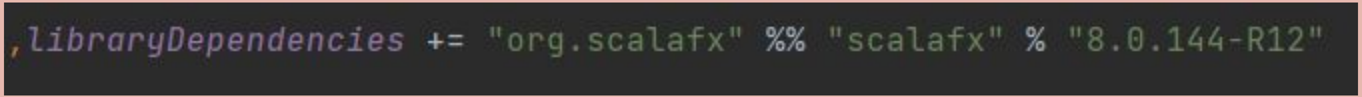


```scala
val mostDangerousDates = crimeDataRDD
  .map {
    row =>
      val date = row(1).toString.split(regex = " ")(0)
      val weapon = row(16)
      (date, weapon)
  }
  .groupByKey() // Group by the date
  .mapValues(_.toList.size) // Map for number of weapons on day
  .sortBy(-1 * _._2) // Sort by weapon count in descending order
  .take(num = 10) // Top 10
```

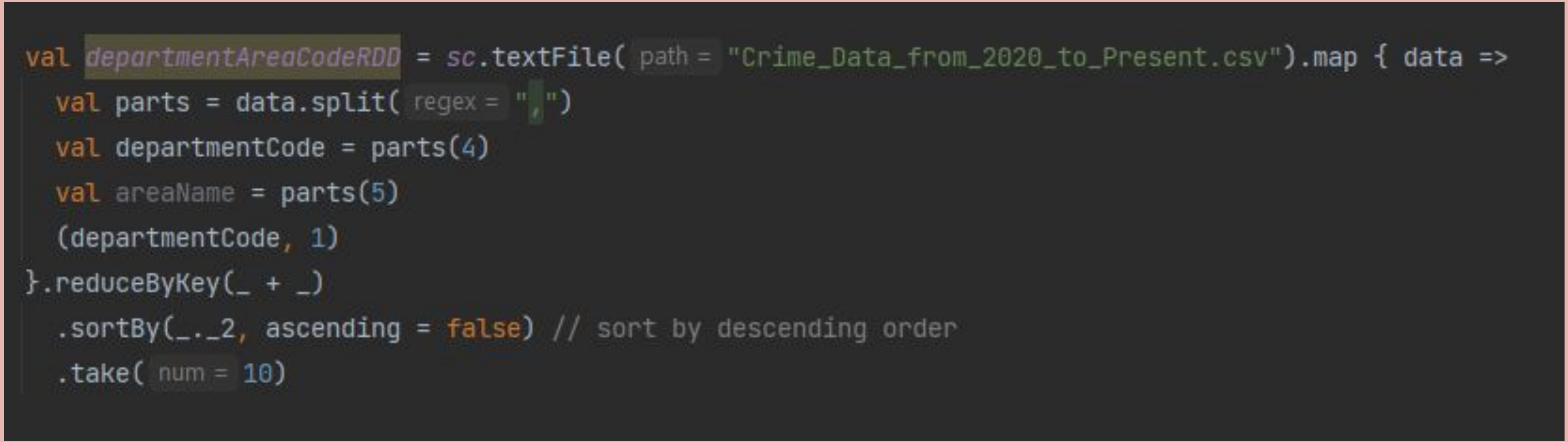


```
(11/01/2021,752)
(10/12/2021,742)
(11/29/2021,727)
(11/08/2021,723)
(09/27/2021,721)
(08/30/2021,720)
(10/18/2021,718)
(07/05/2021,716)
(11/15/2021,712)
(12/07/2021,708)
```

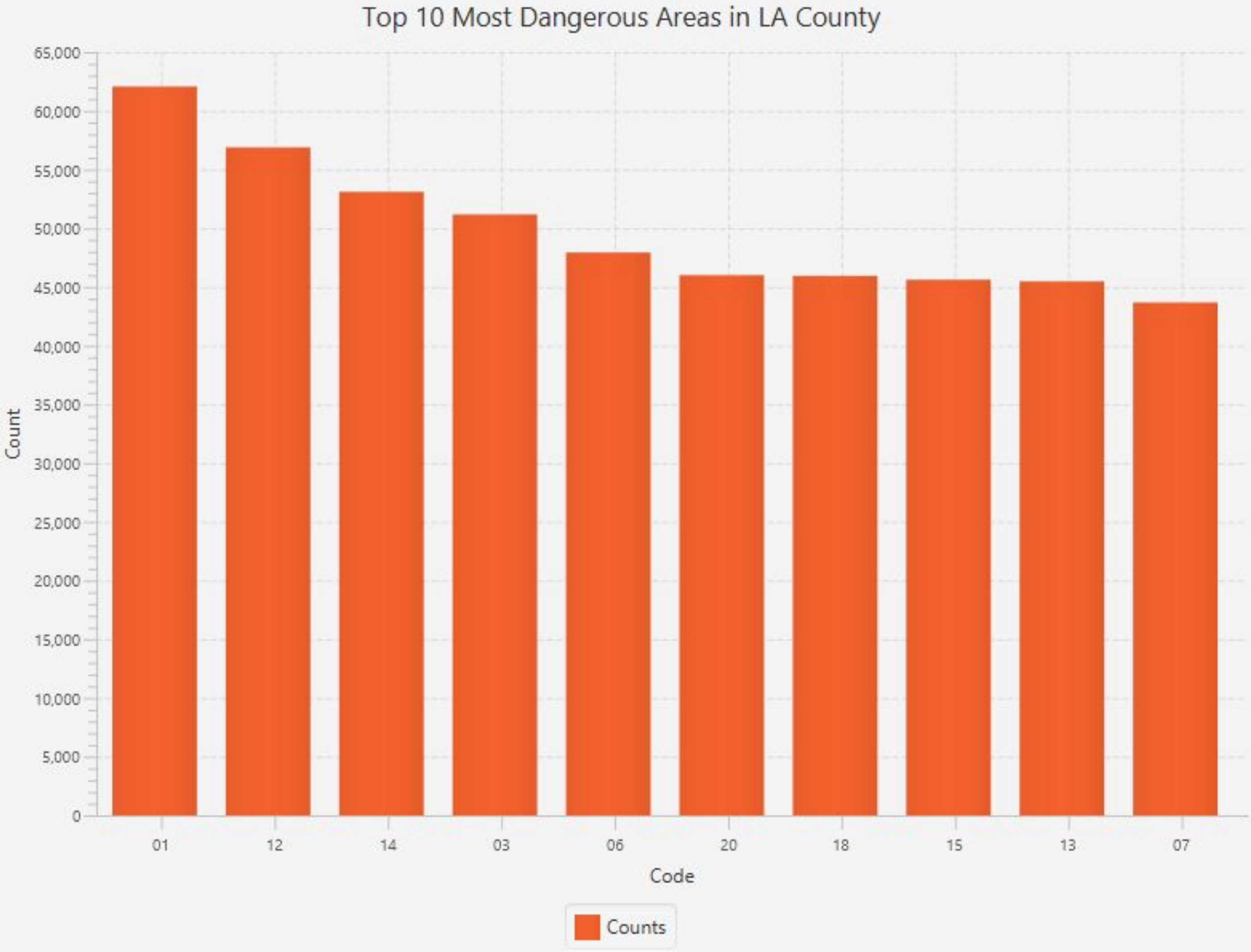
Most Dangerous Areas in LA County

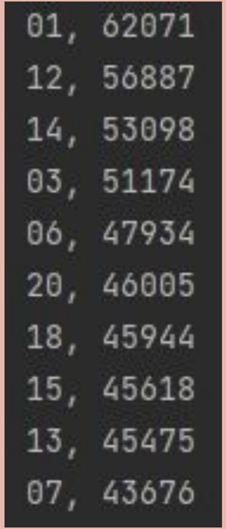


```
,libraryDependencies += "org.scalafx" %% "scalafx" % "8.0.144-R12"
```

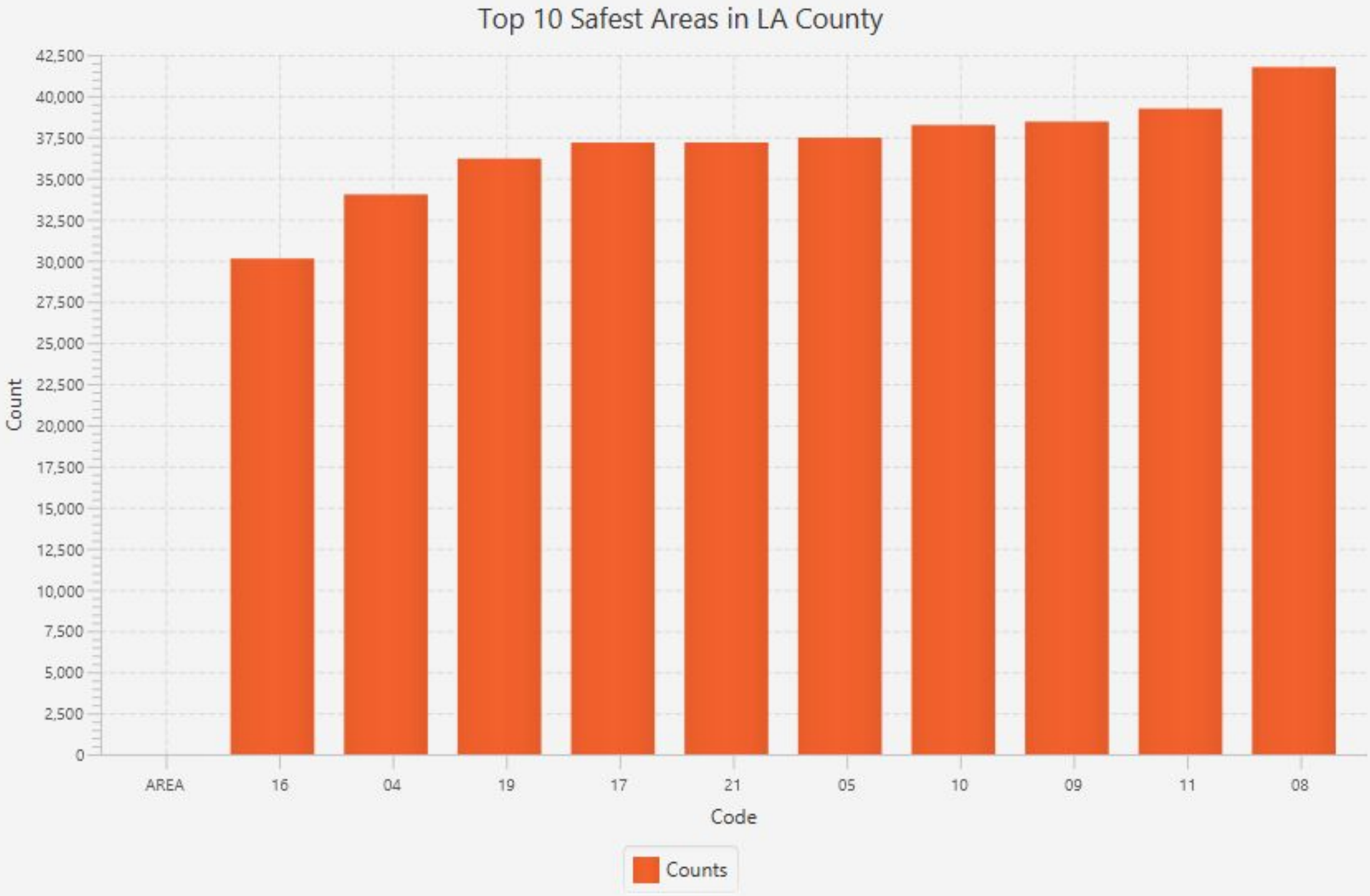


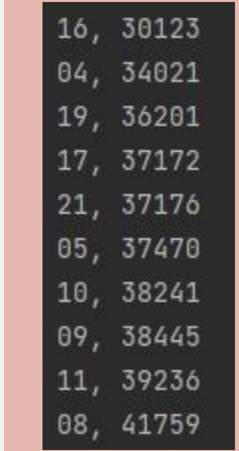
```
val departmentAreaCodeRDD = sc.textFile( path = "Crime_Data_from_2020_to_Present.csv").map { data =>
  val parts = data.split( regex = ",")
  val departmentCode = parts(4)
  val areaName = parts(5)
  (departmentCode, 1)
}.reduceByKey(_ + _)
  .sortBy(_._2, ascending = false) // sort by descending order
  .take( num = 10)
```



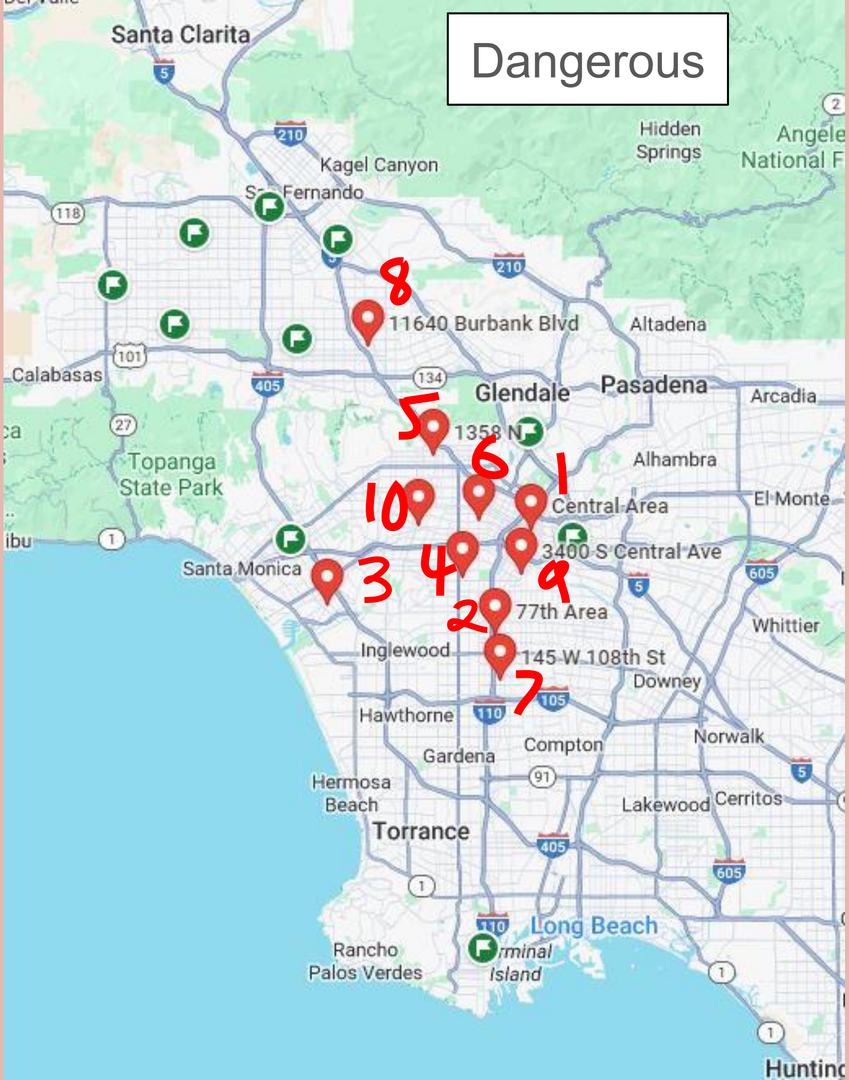


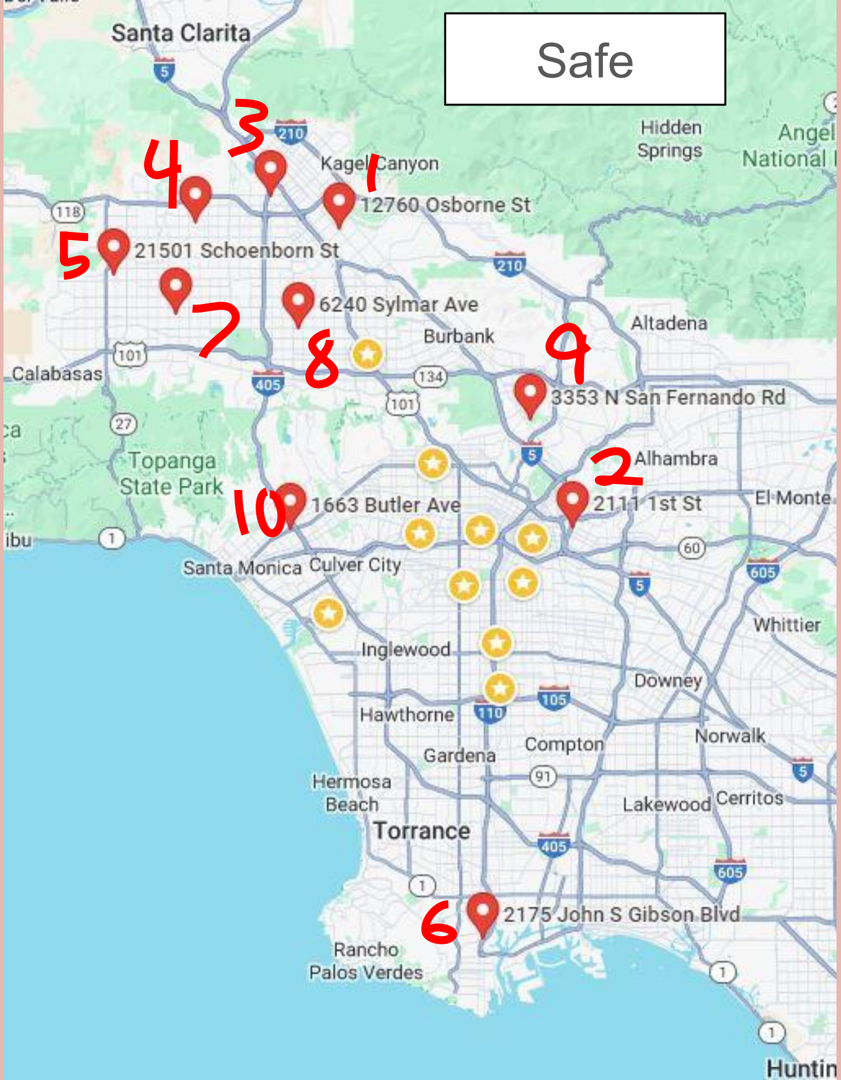
```
01, 62071
12, 56887
14, 53098
03, 51174
06, 47934
20, 46005
18, 45944
15, 45618
13, 45475
07, 43676
```





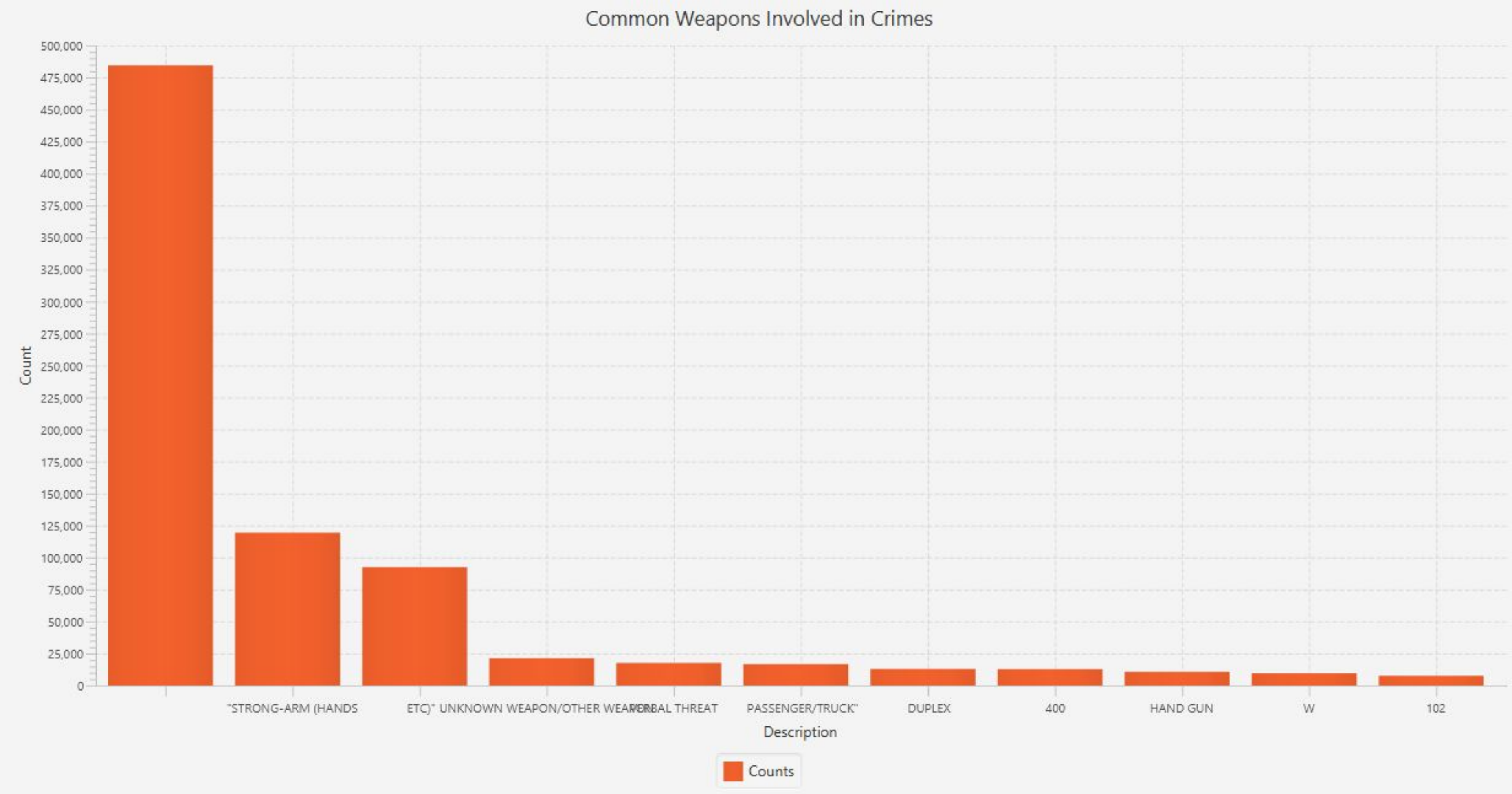
```
16, 30123
04, 34021
19, 36201
17, 37172
21, 37176
05, 37470
10, 38241
09, 38445
11, 39236
08, 41759
```


Dangerous


Safe

Most Common “Weapons” Used - Weapon Desc

1. Most crimes did not involve a weapon: 484,531
2. STRONG-ARM (HANDS): 119,366
3. Unknown WEAPON/OTHER WEAPON: 21,311
4. VERBAL THREAT: 17,648
5. PASSENGER/TRUCK: 16,656
6. HANDGUN: 10,661

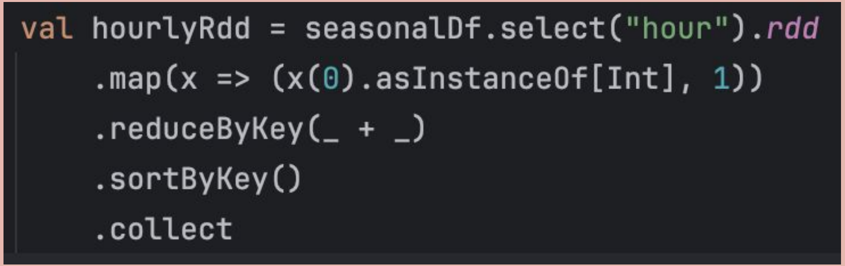


Common Weapons Involved in Crimes

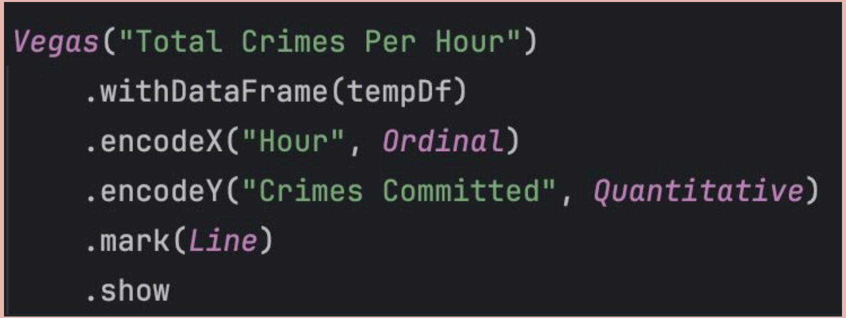
Seasonal Variation in Crimes

Methodology

- Map/Reduce to get counts
- Turn result into a Spark DataFrame and display using Vega-lite



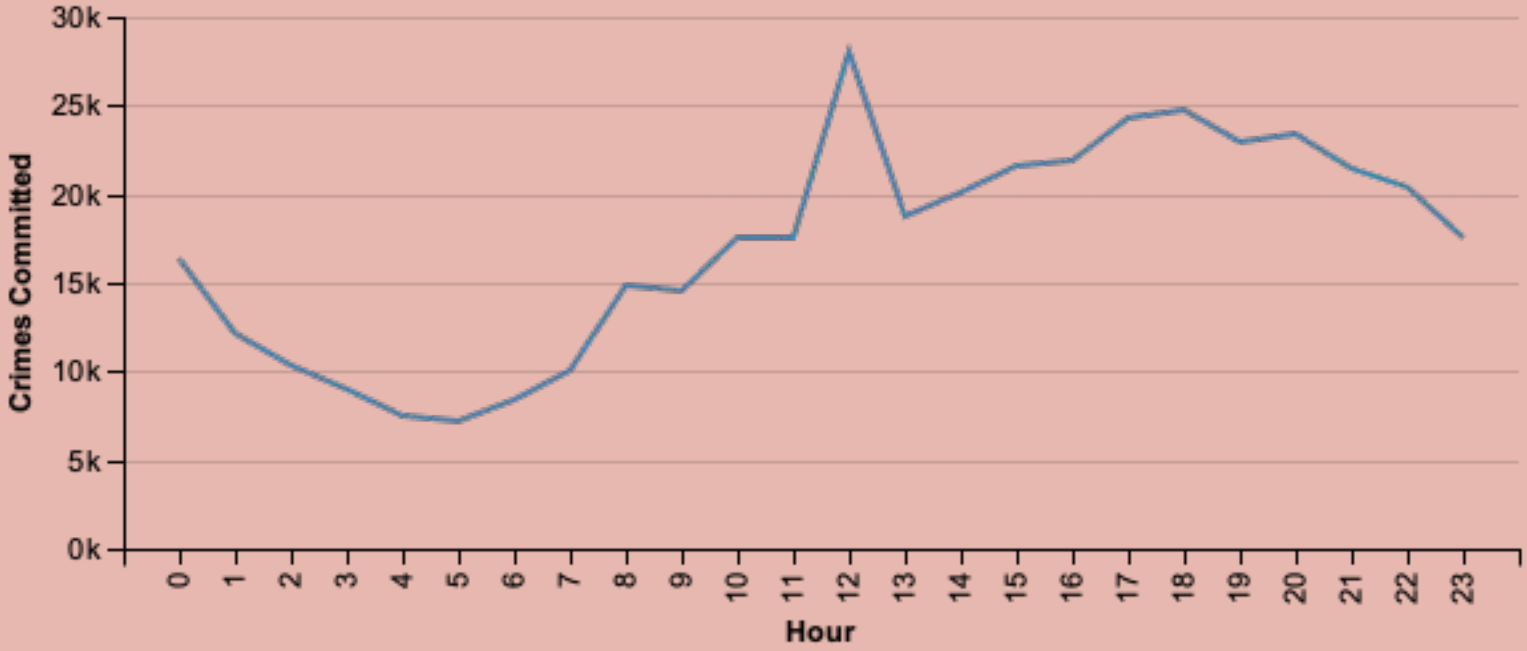
```
val hourlyRdd = seasonalDf.select("hour").rdd
    .map(x => (x(0).asInstanceOf[Int], 1))
    .reduceByKey(_ + _)
    .sortByKey()
    .collect
```



```
Vegas("Total Crimes Per Hour")
    .withDataFrame(tempDf)
    .encodeX("Hour", Ordinal)
    .encodeY("Crimes Committed", Quantitative)
    .mark(Line)
    .show
```

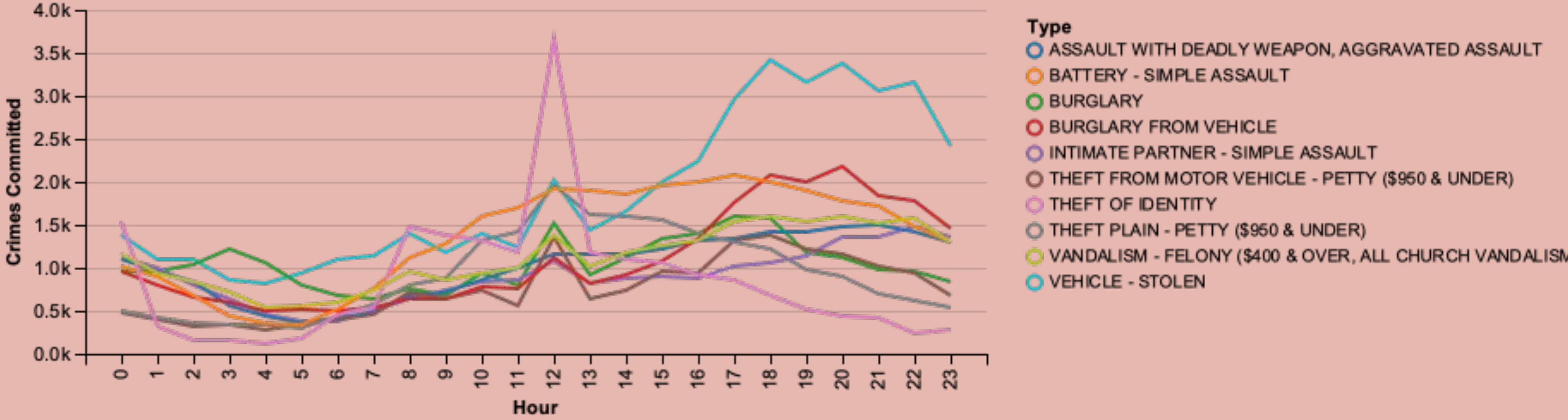
By Day

- LA county uses 12:00pm as a placeholder if a time is missing/unknown, so this value is unreliable.
- Crimes peak at 8-10pm, and fall in the early morning.



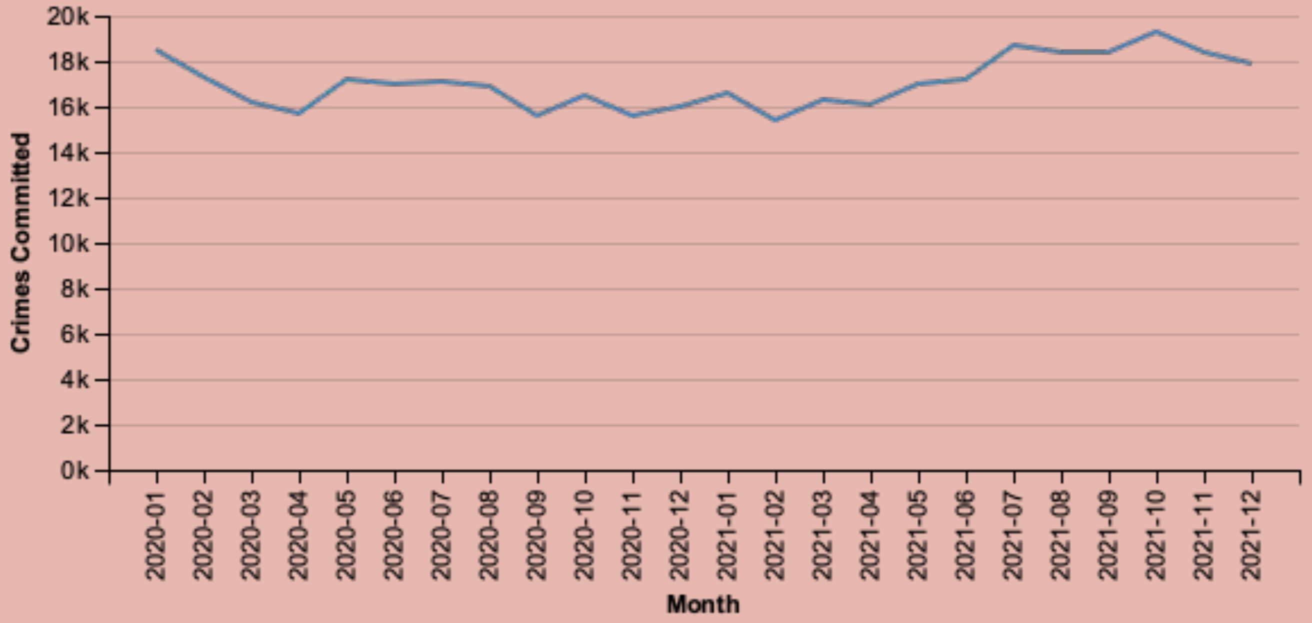
By Day - Observations

- Assaults and vandalism are more common during waking hours, and less common during the night.
- Vehicle theft and burglary spike during the night.



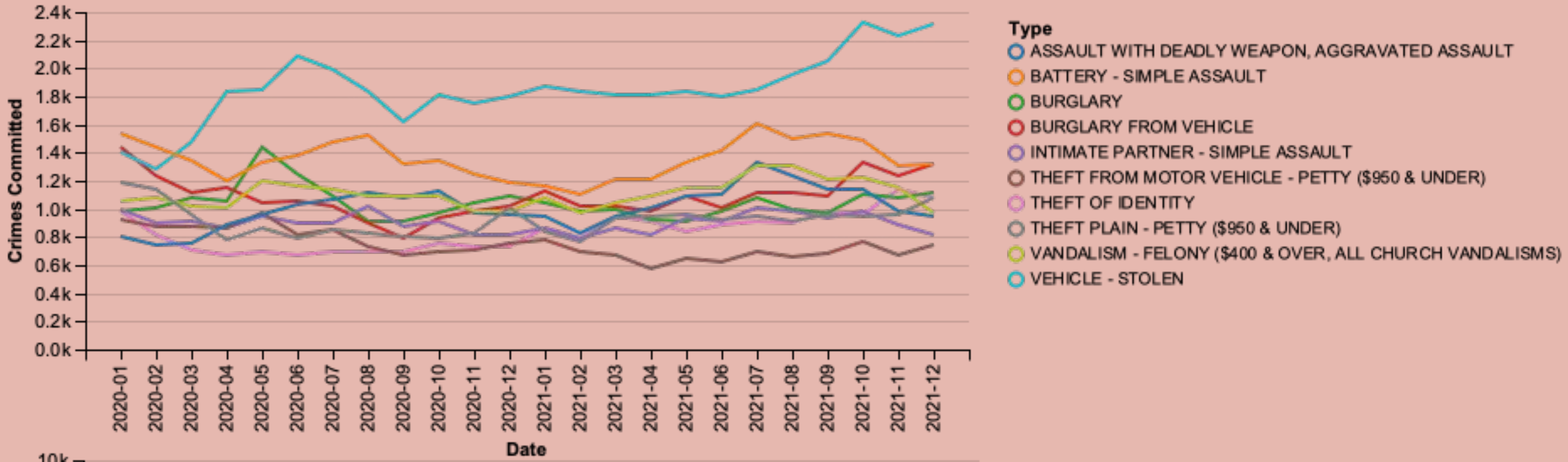
By Month

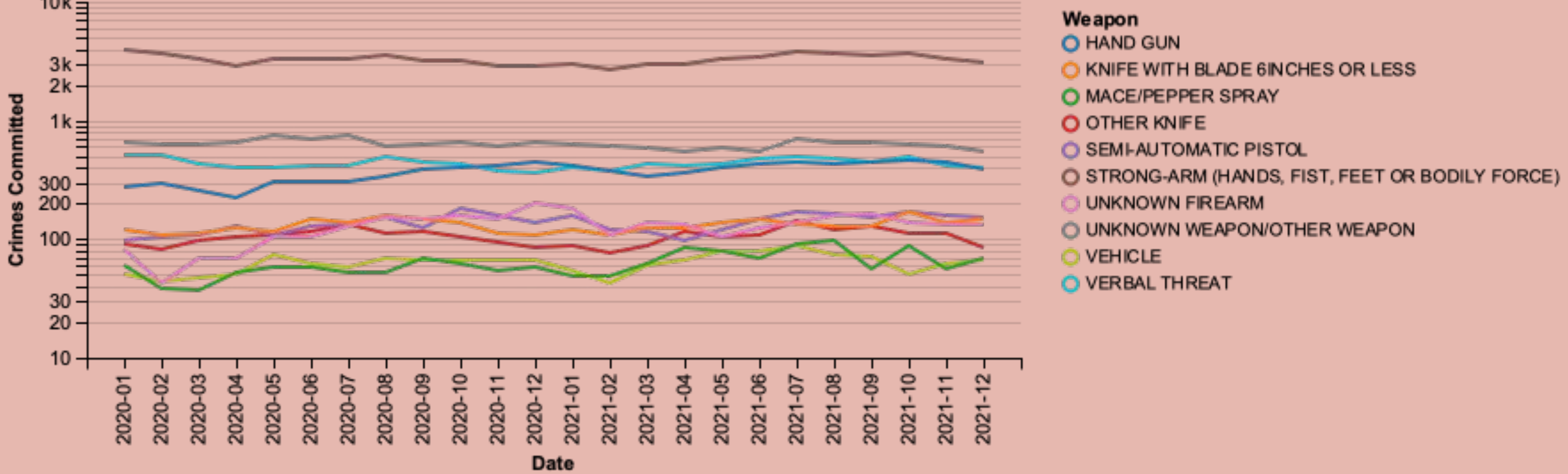
- Crimes show a slight rebound/recovery due to COVID lockdowns, but this could just be noise.

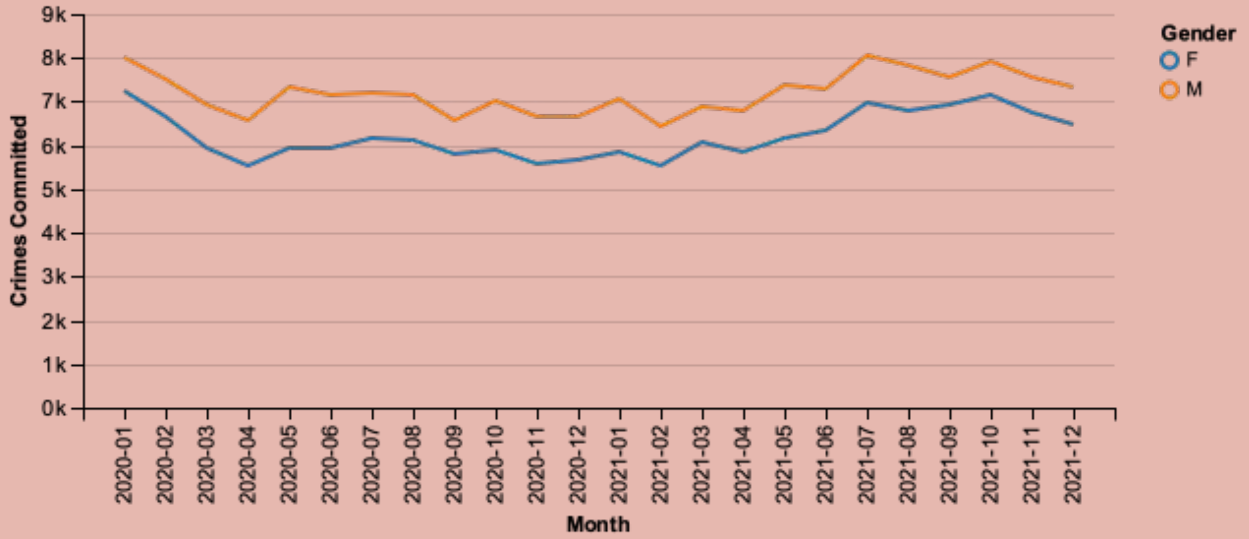


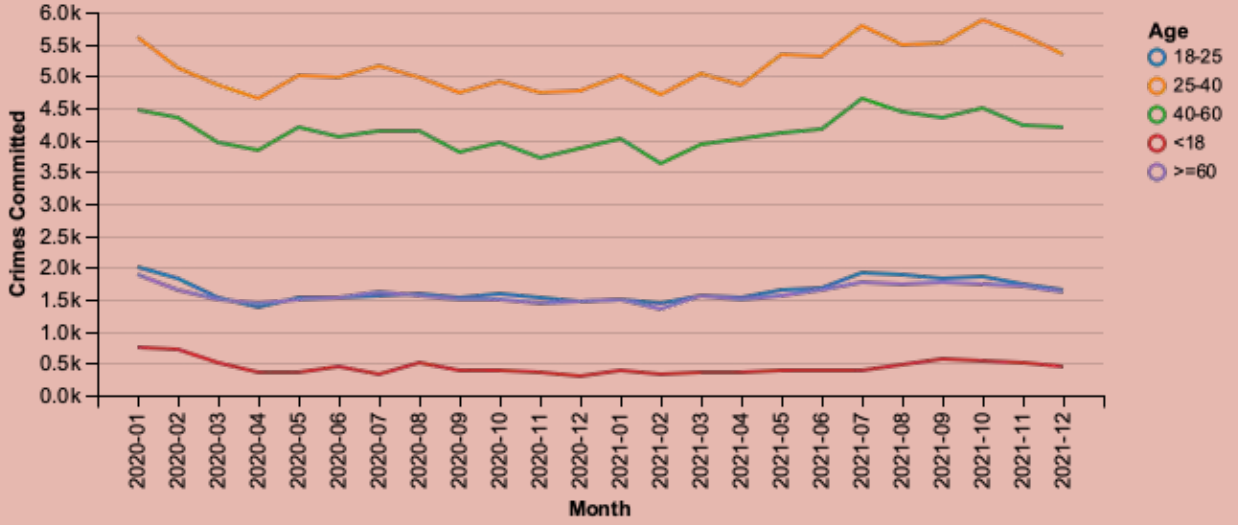
Top 5 Months - Total Reported Crime

2021-10	19316
2021-07	18666
2020-01	18528
2021-08	18384
2021-09	18370


Crimes Committed
2.4k
2.2k
2.0k
1.8k
1.6k
1.4k
1.2k
1.0k
0.8k
0.6k
0.4k
0.2k
0.0k
2020-01
2020-02
2020-03
2020-04
2020-05
2020-06
2020-07
2020-08
2020-09
2020-10
2020-11
2020-12
2021-01
2021-02
2021-03
2021-04
2021-05
2021-06
2021-07
2021-08
2021-09
2021-10
2021-11
2021-12
Date
Type
○ ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT
○ BATTERY - SIMPLE ASSAULT
○ BURGLARY
○ BURGLARY FROM VEHICLE
○ INTIMATE PARTNER - SIMPLE ASSAULT
○ THEFT FROM MOTOR VEHICLE - PETTY ($950 & UNDER)
○ THEFT OF IDENTITY
○ THEFT PLAIN - PETTY ($950 & UNDER)
○ VANDALISM - FELONY ($400 & OVER, ALL CHURCH VANDALISMS)
○ VEHICLE - STOLEN

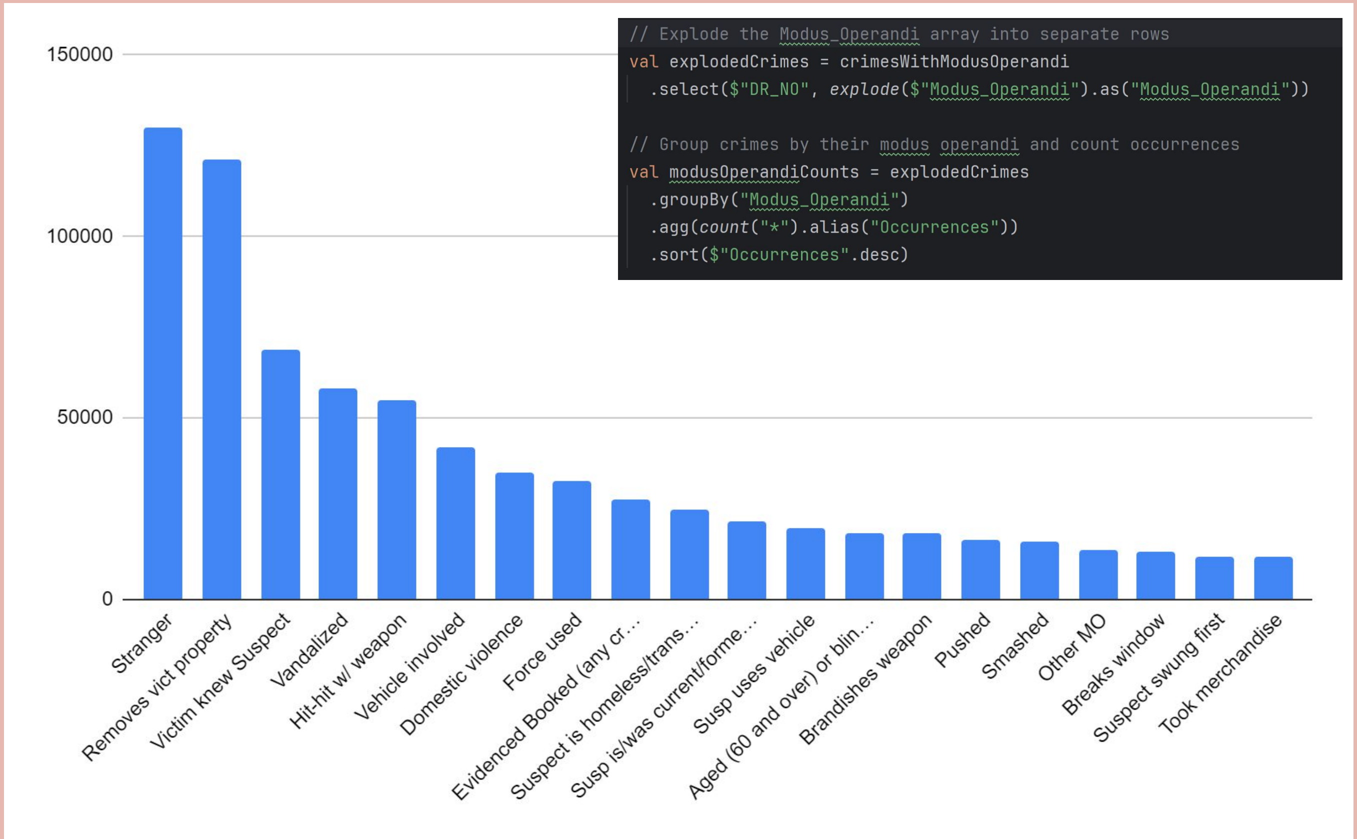

Crimes Committed
10k
3k
2k
1k
300
200
100
30
20
10
2020-01
2020-02
2020-03
2020-04
2020-05
2020-06
2020-07
2020-08
2020-09
2020-10
2020-11
2020-12
2021-01
2021-02
2021-03
2021-04
2021-05
2021-06
2021-07
2021-08
2021-09
2021-10
2021-11
2021-12
Date
Weapon
○ HAND GUN
○ KNIFE WITH BLADE 6INCHES OR LESS
○ MACE/PEPPER SPRAY
○ OTHER KNIFE
○ SEMI-AUTOMATIC PISTOL
○ STRONG-ARM (HANDS, FIST, FEET OR BODILY FORCE)
○ UNKNOWN FIREARM
○ UNKNOWN WEAPON/OTHER WEAPON
○ VEHICLE
○ VERBAL THREAT





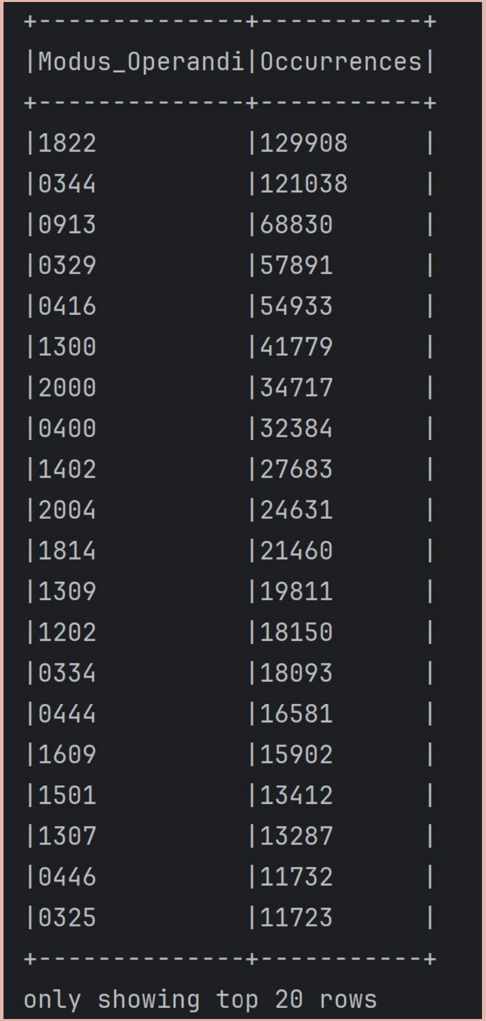
- No significant changes in gender or age group over time.

Most Commonly Reported MOs



```scala
// Explode the Modus_Operandi array into separate rows
val explodedCrimes = crimesWithModusOperandi
  .select($"DR_NO", explode($"Modus_Operandi").as("Modus_Operandi"))

// Group crimes by their modus operandi and count occurrences
val modusOperandiCounts = explodedCrimes
  .groupBy("Modus_Operandi")
  .agg(count("*").alias("Occurrences"))
  .sort($"Occurrences".desc)
```



```
+--------------+-----------+
|Modus_Operandi|Occurrences|
+--------------+-----------+
|1822          |129908     |
|0344          |121038     |
|0913          |68830      |
|0329          |57891      |
|0416          |54933      |
|1300          |41779      |
|2000          |34717      |
|0400          |32384      |
|1402          |27683      |
|2004          |24631      |
|1814          |21460      |
|1309          |19811      |
|1202          |18150      |
|0334          |18093      |
|0444          |16581      |
|1609          |15902      |
|1501          |13412      |
|1307          |13287      |
|0446          |11732      |
|0325          |11723      |
+--------------+-----------+
only showing top 20 rows
```

Conclusions

Obstacles

- Problems with missing or misleading default data.
- Not as much documentation available for scala when compared to Java, Python.

What We Learned

- We used scala/spark to look at trends in crime location, demographics, and methods.
- The most common crimes are vehicle and burglary related.
- Police stations that handle more crimes are in more populous areas.
- Theft is much more common during the night, little variation in monthly patterns.
- Crime habits commonly involve anonymity and erratic behaviors