

HADOOP TOOL USED: Impala

ANALYTICS-1 :

Analysing the number of fire calls & building violations for areas based on their zipcodes and over years and get a pattern [\[Tables: Building violations, Fire calls\]](#)

// Analysing total calls and violation over area

```
impala-shell -i compute-1-1:21000 -q 'select zipcode, count(*) from
pa1373.fire_incidents as fi join pa1373.firecalls as fc on fi.incident_no =
fc.incident_number where year >= "2008" and year<= "2018" group by zipcode order by
zipcode;' -B -o analytics11.csv
```

```
impala-shell -i compute-1-1:21000 -q 'select zipcode, count(*) from pa1373.viola where
year >= "2008" and year<= "2018" group by zipcode order by zipcode;' -B -o
analytics12.csv
```

ANALYTICS-2_:

Analysing the number of violations over different areas and years according to nov_category [\[Building violations\]](#)

// getting top 5 areas(zipcodes) where max violations are reported

```
select zipcode,count(*) from building_violations group by zipcode order by count(*) desc
limit 5;
```

// for each area: find the change in the count of violations of different types (nov_category description) over the past 10 years

```
impala-shell -i compute-1-1:21000 -q 'Select * from (Select
nov_category_description,year,count(*) from (Select * from sg5783.building_violations
where zipcode=94110) sub1 where year> "2008" and year< "2018" group by
sub1.nov_category_description, sub1.year) sub2 where
sub2.nov_category_description="fire section" or sub2.nov_category_description="smoke
detection section" or sub2.nov_category_description="building section" or
sub2.nov_category_description="plumbing and electrical section" or
sub2.nov_category_description="interior surfaces section" order by
sub2.nov_category_description asc, sub2.year asc' -B -o analytics21.csv;
```

```
impala-shell -i compute-1-1:21000 -q 'Select * from (Select  
nov_category_description,year,count(*) from (Select * from sg5783.building_violations  
where zipcode=94109) sub1 where year> "2008" and year< "2018" group by  
sub1.nov_category_description, sub1.year) sub2 where  
sub2.nov_category_description="fire section" or sub2.nov_category_description="smoke  
detection section" or sub2.nov_category_description="building section" or  
sub2.nov_category_description="plumbing and electrical section" or  
sub2.nov_category_description="interior surfaces section" order by  
sub2.nov_category_description asc, sub2.year asc' -B -o analytics22.csv;
```

```
impala-shell -i compute-1-1:21000 -q 'Select * from (Select  
nov_category_description,year,count(*) from (Select * from sg5783.building_violations  
where zipcode=94117) sub1 where year> "2008" and year< "2018" group by  
sub1.nov_category_description, sub1.year) sub2 where  
sub2.nov_category_description="fire section" or sub2.nov_category_description="smoke  
detection section" or sub2.nov_category_description="building section" or  
sub2.nov_category_description="plumbing and electrical section" or  
sub2.nov_category_description="interior surfaces section" order by  
sub2.nov_category_description asc, sub2.year asc' -B -o analytics23.csv;
```

```
impala-shell -i compute-1-1:21000 -q 'Select * from (Select  
nov_category_description,year,count(*) from (Select * from sg5783.building_violations  
where zipcode=94102) sub1 where year> "2008" and year< "2018" group by  
sub1.nov_category_description, sub1.year) sub2 where  
sub2.nov_category_description="fire section" or sub2.nov_category_description="smoke  
detection section" or sub2.nov_category_description="building section" or  
sub2.nov_category_description="plumbing and electrical section" or  
sub2.nov_category_description="interior surfaces section" order by  
sub2.nov_category_description asc, sub2.year asc' -B -o analytics24.csv;
```

```
impala-shell -i compute-1-1:21000 -q 'Select * from (Select  
nov_category_description,year,count(*) from (Select * from sg5783.building_violations  
where zipcode=94103) sub1 where year> "2008" and year< "2018" group by  
sub1.nov_category_description, sub1.year) sub2 where  
sub2.nov_category_description="fire section" or sub2.nov_category_description="smoke  
detection section" or sub2.nov_category_description="building section" or  
sub2.nov_category_description="plumbing and electrical section" or
```

```
sub2.nov_category_description="interior surfaces section" order by  
sub2.nov_category_description asc, sub2.year asc' -B -o analytics25.csv;
```

ANALYTICS-3_:

Calculate false alarms over the years and compare it with violations having nov category fire, smoke , plumbing and electrical over the years for different neighbourhoods [\[Fire_calls_service, fire_incidents, Building\]](#)

```
impala-shell -i compute-1-1:21000 -q 'select year, count(*) from pa1373.firecalls as fc  
join pa1373.fire_incidents as fi on fc.incident_number = fi.incident_no where  
instr(primary_situation, "false") > 0 OR instr(primary_situation, "unintentional")>0  
group by year order by year desc;' -B -o analytics31.csv
```

```
impala-shell -i compute-1-1:21000 -q 'select zipcode, count(*) from pa1373.firecalls as  
fc join pa1373.fire_incidents as fi on fc.incident_number = fi.incident_no where  
instr(primary_situation, "false") > 0 OR instr(primary_situation, "unintentional")>0  
group by zipcode order by zipcode desc;' -B -o analytics32.csv
```

ANALYTICS-4_:

Analysing call_type over the years over different areas according to neighbourhood boundaries [\[Fire_calls_service\]](#)

// getting top 5 areas(neighborhood_analysis_boundaries) where max calls are reported

```
select neighborhood_analysis_boundaries,count(*) from fire_calls_service group by  
neighborhood_analysis_boundaries order by count(*) desc limit 5;
```

// for each neighbourhood area: find the change in the number of calls of different types
(call_type) over the past 10 years

```
impala-shell -i compute-1-1:21000 -q "select * from (Select  
neighborhoodanalysisboundaries, call_type, year, count(*) from (Select * from  
pa1373.firecalls) as sub1 where sub1.year >='2008' and sub1.year <= '2018' group by  
sub1.call_type, sub1.year, neighborhoodanalysisboundaries) as sub2 where
```

```
sub2.call_type = 'Citizen Assist / Service Call' or call_type = 'Electrical Hazard' or  
call_type = 'Elevator / Escalator Rescue' or call_type = 'Gas Leak (Natural and LP  
Gases)' or call_type = 'Structure Fire' order by call_type;" -B -o analytics4-all.csv
```

```
impala-shell -i compute-1-1:21000 -q "select * from (Select call_type, year, count(*)  
from (Select * from pa1373.firecalls where neighborhoodsanalysisboundaries =  
'Tenderloin') as sub1 where sub1.year >='2008' and sub1.year <= '2018' group by  
sub1.call_type, sub1.year) as sub2 where sub2.call_type = 'Citizen Assist / Service Call'  
or call_type = 'Electrical Hazard' or call_type = 'Elevator / Escalator Rescue' or call_type  
= 'Gas Leak (Natural and LP Gases)' or call_type = 'Structure Fire' order by call_type;" -  
B -o analytics41.csv
```

```
impala-shell -i compute-1-1:21000 -q "select * from (Select call_type, year, count(*)  
from (Select * from pa1373.firecalls where neighborhoodsanalysisboundaries =  
'Mission') as sub1 where sub1.year >='2008' and sub1.year <= '2018' group by  
sub1.call_type, sub1.year) as sub2 where sub2.call_type = 'Citizen Assist / Service Call'  
or call_type = 'Electrical Hazard' or call_type = 'Elevator / Escalator Rescue' or call_type  
= 'Gas Leak (Natural and LP Gases)' or call_type = 'Structure Fire' order by call_type;" -  
B -o analytics42.csv
```

```
impala-shell -i compute-1-1:21000 -q "select * from (Select call_type, year, count(*)  
from (Select * from pa1373.firecalls where neighborhoodsanalysisboundaries =  
'Financial District/South Beach') as sub1 where sub1.year >= '2008' and sub1.year <=  
'2018' group by sub1.call_type, sub1.year) as sub2 where sub2.call_type = 'Citizen Assist  
/ Service Call' or call_type = 'Electrical Hazard' or call_type = 'Elevator / Escalator  
Rescue' or call_type = 'Gas Leak (Natural and LP Gases)' or call_type = 'Structure Fire'  
order by call_type;" -B -o analytics43.csv
```

```
impala-shell -i compute-1-1:21000 -q "select * from (Select call_type, year, count(*)  
from (Select * from pa1373.firecalls where neighborhoodsanalysisboundaries =  
'Bayview Hunters Point') as sub1 where sub1.year >= '2008' and sub1.year <= '2018'  
group by sub1.call_type, sub1.year) as sub2 where sub2.call_type = 'Citizen Assist /  
Service Call' or call_type = 'Electrical Hazard' or call_type = 'Elevator / Escalator Rescue'  
or call_type = 'Gas Leak (Natural and LP Gases)' or call_type = 'Structure Fire' order by  
call_type;" -B -o analytics44.csv
```

```
impala-shell -i compute-1-1:21000 -q "select * from (Select call_type, year, count(*)  
from (Select * from pa1373.firecalls where neighborhoodsanalysisboundaries = 'South  
of Market') as sub1 where sub1.year >= '2008' and sub1.year <= '2018' group by  
sub1.call_type, sub1.year) as sub2 where sub2.call_type = 'Citizen Assist / Service Call'  
or call_type = 'Electrical Hazard' or call_type = 'Elevator / Escalator Rescue' or call_type  
= 'Gas Leak (Natural and LP Gases)' or call_type = 'Structure Fire' order by call_type;" -  
B -o analytics45.csv
```

