**SQL Code summary – Seattle Vehicles Collisions Analysis**

Team 2: Bryan Wang, Hao Tung, Jess Lu, Neelabja Gayen, Pallavi Khabale

#for the use of this table, please refer to all words in Red to understand what the section is about.

# Below refers to Time of Day Analysis

**select** Year, Month, *s*.Weekday, *s*.TIMEOFDAY , **count**(\*) **as** *"TotalCase"*

**from** sdotcol *s*

**where** *s*.TIMEOFDAY <> ''

**group** **by** Year, Month, *s*.Weekday, *s*.TIMEOFDAY

**order** **by** **count**(\*) **desc** **limit** 10;

**SELECT**

*s*.TIMEOFDAY,

**COUNT**(\*) **AS** *"TotalCase"*,

**ROUND**(**COUNT**(\*) \* 100.0 / **SUM**(**COUNT**(\*)) **OVER** (), 0) **AS** *"Percentage"*

**FROM** sdotcol *s*

**WHERE** *s*.TIMEOFDAY <> ''

**GROUP** **BY** *s*.TIMEOFDAY

**order** **by** *Percentage* **desc** ;

# Below refers to Severity & Speeding

--C1-1 Total case by severity

**select**

SEVERITYCODE

, **count**(SEVERITYCODE)

**from** sdotcol *s*

**where** severitycode != ('0' **or** **''**)

**group** **by** *s*.SEVERITYCODE ;

--C 2-1 Collision count by year

**select** Year

, **count**(\*)

**from** sdotcol *s*

**group** **by** Year

--C 2-2 collision count by month (can put dif year on same month bar)

**select** Month

, **count**(\*)

**from** sdotcol *s*

**group** **by** Month

--C 2-3 collision by time, can sum over years.

**select** *s*.TIMEOFDAY

, **count**(\*)

**from** sdotcol *s*

**group** **by** TIMEOFDAY

--C 2-4 Severity count based on dif time periods

**select**

*s*.TIMEOFDAY

, *s*.SEVERITYCODE

, **count**(\*)

**from** sdotcol *s*

**group** **by** SEVERITYCODE, TIMEOFDAY

--C 3-2 Severity distribution from speeding

**SELECT**

SEVERITYCODE,

**COUNT**(\*) **AS** *count\_of\_severity*,

(**COUNT**(\*) \* 100.0 / (**SELECT** **COUNT**(\*)

**FROM** sdotcol

**WHERE** Speeding = 'Y')) **AS** *percentage*

**FROM** sdotcol

**WHERE** Speeding = 'Y'

**GROUP** **BY** SEVERITYCODE;

--C 3-3 Severity distribution from non-speeding

**SELECT**

SEVERITYCODE,

**COUNT**(\*) **AS** *count\_of\_severity*,

(**COUNT**(\*) \* 100.0 / (**SELECT** **COUNT**(\*)

**FROM** sdotcol

**WHERE** Speeding = 'N')) **AS** *percentage*

**FROM** sdotcol

**WHERE** Speeding = 'N'

**GROUP** **BY** SEVERITYCODE;

--C 3-4 % of accidents occurred from speeding by time periods

**select**

timeofday,

**count**(\*) **as** *count\_of\_severity*,

(**COUNT**(\*) \* 100.0 / (**SELECT** **COUNT**(\*)

**FROM** sdotcol

**WHERE** Speeding = 'Y')) **AS** *percentage*

**FROM** sdotcol

**where** speeding = 'Y'

**group** **by** timeofday;

--C 3-5 count of each type of collision with speeding

**select** sdot\_coldesc, **count**(**case** **when** speeding = 'Y' **then** 1 **end**) **as** *SpdYes*

**FROM** sdotcol

**group** **by** sdot\_coldesc

**order** **by** *SpdYes* **desc**

**limit** 10;

--C 3-6 Count of speeding collision locations

**select** location, **count**(**case** **when** speeding = 'Y' **then** 1 **end**) **as** *SpdYes*

**FROM** sdotcol

**group** **by** location

**order** **by** *SpdYes* **desc**

**limit** 10;

# Below refers to inattentiveness

--C 4-1 % of collision cause by not paying attention by time periods

**select**

TIMEOFDAY

, **count**(Inattentionind)

**from** sdotcol *s*

**where** Inattentionind = 'Y'

**group** **by** Inattentionind, TIMEOFDAY;

--C 4-2 count of Collision type (inattentive)

**select**

*s*.COLLISIONTYPE

, **count**(Inattentionind) **as** *count*

**from** sdotcol *s*

**where** Inattentionind = 'Y'

**group** **by** COLLISIONTYPE

# Below refers to collision type

--C 6-1 Table between time periods and collision type

**select**

*s*.TIMEOFDAY

, *s*.COLLISIONTYPE

, **Count**(\*)

**from** sdotcol *s*

**group** **by** COLLISIONTYPE, TIMEOFDAY

--C 6-2 Pie chart compare left turn and right turn collision

**select**

*s*.COLLISIONTYPE

, **Count**(\*)

, **ROUND**(**COUNT**(\*) \* 100.0 / **SUM**(**COUNT**(\*)) **OVER** (), 2) **AS** *percentage*

**from** sdotcol *s*

**where** COLLISIONTYPE ='Left Turn' **or** COLLISIONTYPE ='Right Turn'

**group** **by** COLLISIONTYPE

-- C 6-3 same as 6-1

--C 6-4 Count of every specific collision type

**select**

collisiontype

, **count**(collisiontype) **as** *Col\_Count*

, **ROUND**( (**COUNT**(collisiontype) \* 100.0) / **SUM**(**COUNT**(collisiontype)) **OVER** (),2 ) **AS** *Col\_Per*

**from** sdotcol *s*

**where** collisiontype <> ''

**group** **by** collisiontype

**order** **by** *Col\_Count* **desc** ;

# Below refers to Collision group by address

--C 7-1 every Collision severity’s % out of total collisions group by address type (ex. Alley, block, intersection)

**select**

*s*.ADDRTYPE

, **sum**(*s*.INJURIES) **as** *INJURIES*

, **sum**(*s*.SERIOUSINJURIES) **as** *SERIOUSINJURIES*

, **sum**(*s*.FATALITIES) **as** *FATALITIES*

, (**SUM**(*s*.INJURIES) + **SUM**(*s*.SERIOUSINJURIES) + **SUM**(*s*.FATALITIES)) **AS** *TOTAL\_INJURIES*

**from** sdotcol *s*

**group** **by** *s*.ADDRTYPE

--C 7-2 every Collision severity’s count group by address type (ex. Alley, block, intersection)

**select**

*s*.ADDRTYPE

, (*s*.SEVERITYCODE ) **as** *SEVERITY*

, **sum**(*s*.INJURIES) **as** *INJURIES*

, **sum**(*s*.SERIOUSINJURIES) **as** *SERIOUSINJURIES*

, **sum**(*s*.FATALITIES) **as** *FATALITIES*

, (**SUM**(*s*.INJURIES) + **SUM**(*s*.SERIOUSINJURIES) + **SUM**(*s*.FATALITIES)) **AS** *TOTAL\_INJURIES*

**from** sdotcol *s*

**group** **by** *s*.ADDRTYPE, *s*.SEVERITYCODE

--C 7-3 collision count by locations

**select**

*s*.ADDRTYPE

, (*s*.COLLISIONTYPE ) **as** *COLLISIONTYPE*

, (**SUM**(*s*.INJURIES) + **SUM**(*s*.SERIOUSINJURIES) + **SUM**(*s*.FATALITIES)) **AS** *TOTAL\_INJURIES*

**from** sdotcol *s*

**group** **by** *s*.ADDRTYPE, *s*.COLLISIONTYPE;

--C 7-4 collision count by address

**select**

intkey

, location

, **count**(intkey) **as** *count*

, **sum**(*s*.INJURIES) **as** *INJURIES*

, **sum**(*s*.SERIOUSINJURIES) **as** *SERIOUSINJURIES*

, **sum**(*s*.FATALITIES) **as** *FATALITIES*

, (**SUM**(*s*.INJURIES) + **SUM**(*s*.SERIOUSINJURIES) + **SUM**(*s*.FATALITIES)) **AS** *TOTAL\_INJURIES*

**from** sdotcol *s*

--where trim(intkey) <>''

**group** **by** intkey, location

**having** *TOTAL\_INJURIES* > **count**(intkey)

**order** **by** fatalities **desc** , *count* **desc**;

--A 7-5-1 Special case - a tragedy caused 5 killed and 119 injured

**select**

sdotcol.INCDATE

, intkey

, location

, injuries

,SERIOUSINJURIES

,FATALITIES

**from** sdotcol

**where** location = 'AURORA BR BETWEEN RAYE ST AND BRIDGE WAY N'

**order** **by** FATALITIES **desc**

# Below refers to Road Condition Analysis

--C 8-1 collision count by road conditions (dry, wet, ice…)

**SELECT**

roadcond,

**COUNT**(\*) **AS** *"Case"*,

**Round** ((**COUNT**(\*) \* 100.0) / **SUM**(**COUNT**(\*)) **OVER** () , 2) **AS** *"Percentage"*

**FROM** sdotcol *s*

**WHERE** roadcond <> ''

**GROUP** **BY** roadcond

**order** **by** *"Case"* **desc**;

--C 8-2 collision count by light conditions (daylight, dusk…)

**SELECT**

*s*.LIGHTCOND ,

**COUNT**(\*) **AS** *"Case"*,

**Round** ((**COUNT**(\*) \* 100.0) / **SUM**(**COUNT**(\*)) **OVER** () , 2) **AS** *"Percentage"*

**FROM** sdotcol *s*

**WHERE** LIGHTCOND <> ''

**GROUP** **BY** LIGHTCOND

**order** **by** *"Case"* **desc**;

--C 8-3 Severity % by dif road condition

**SELECT**

*s*.SEVERITYCODE ,

roadcond,

**COUNT**(\*) **AS** *"Case"*

-- Round ((COUNT(\*) \* 100.0) / SUM(COUNT(\*)) OVER () , 2) AS "Percentage"

**FROM** sdotcol *s*

**WHERE** roadcond <> '' **and** SEVERITYCODE <> ('' **or** **'Unknown'**)

**GROUP** **BY** SEVERITYCODE, roadcond

**order** **by** SEVERITYCODE , *"Case"* **desc**;

--C 8-4 Severity % by dif light condition

**SELECT**

*s*.SEVERITYCODE ,

LIGHTCOND,

**COUNT**(\*) **AS** *"Case"*

-- Round ((COUNT(\*) \* 100.0) / SUM(COUNT(\*)) OVER () , 2) AS "Percentage"

**FROM** sdotcol *s*

**WHERE** LIGHTCOND <> '' **and** SEVERITYCODE <> ('' **or** **'Unknown'**)

**GROUP** **BY** SEVERITYCODE, LIGHTCOND

**order** **by** SEVERITYCODE , *"Case"* **desc**;

# Below refers to COVID impact analysis

**SELECT**

'Before 2020' **AS** *Year*,

**ROUND**(**AVG**(*year\_count* ), 2) **AS** *"Case per Year"*

**FROM** (

**SELECT** Year, **COUNT**(\*) **AS** *year\_count*

**FROM** sdotcol

**WHERE** Year < '2020' **and** Year <> '2003'

**GROUP** **BY** Year

) **AS** *yearly\_counts*

**Union** **all**

**select** '2020' **as** *Year*,

**Count**(\*) **as** *"Case per Year"*

**from** sdotcol

**Where** Year = '2020'