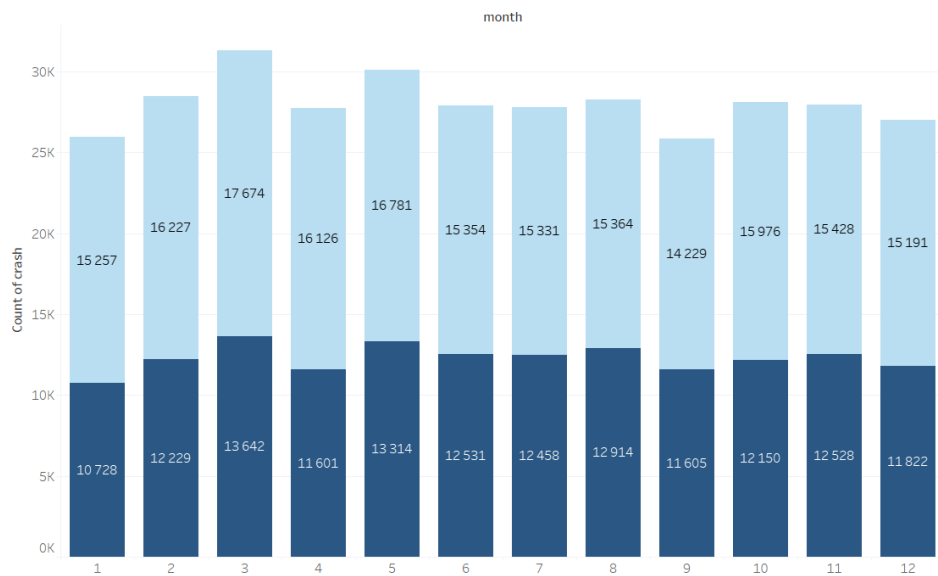


Wykonanie analizy:

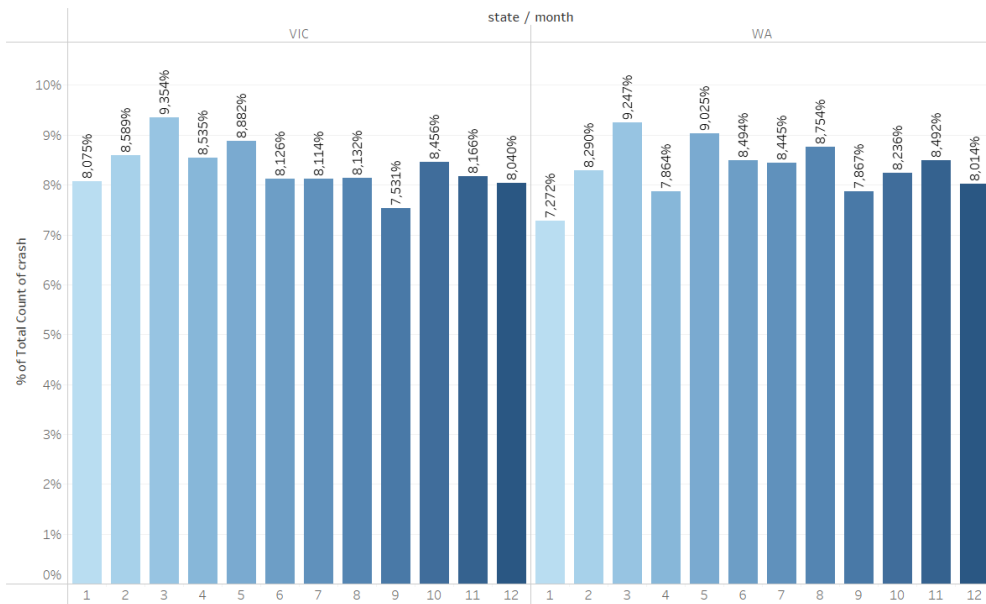
1. Ile było wypadków w zależności od miesiąca oraz stanu("state") w latach 2006-2019.

State	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
VIC	15,257	16,227	17,674	16,126	16,781	15,354	15,331	15,364	14,229	15,976	15,428	15,191
WA	10,728	12,229	13,642	11,601	13,314	12,531	12,458	12,914	11,605	12,150	12,528	11,822

Ile było wypadków w zależności od miesiąca oraz stanu("state") w latach 2006-2019.



Ile było wypadków w zależności od miesiąca oraz stanu("state") w latach 2006-2019.



Obserwacje:

Na pierwszym i drugim wykresie możemy zobaczyć że bezwzględnie najwięcej wypadków było w stanie „VIC”, ale jeśli popatrzymy na ostatni wykres gdzie mamy odsetki dla każdego stanu osobno – możemy zobaczyć że procentowo w różne miesiące mamy większe odsetki w stanie „WA” np.:

Do maja mamy więcej wypadków w stanie „VIC” ale od maja do października mamy odsetkowo więcej wypadków same w stanie „WA”

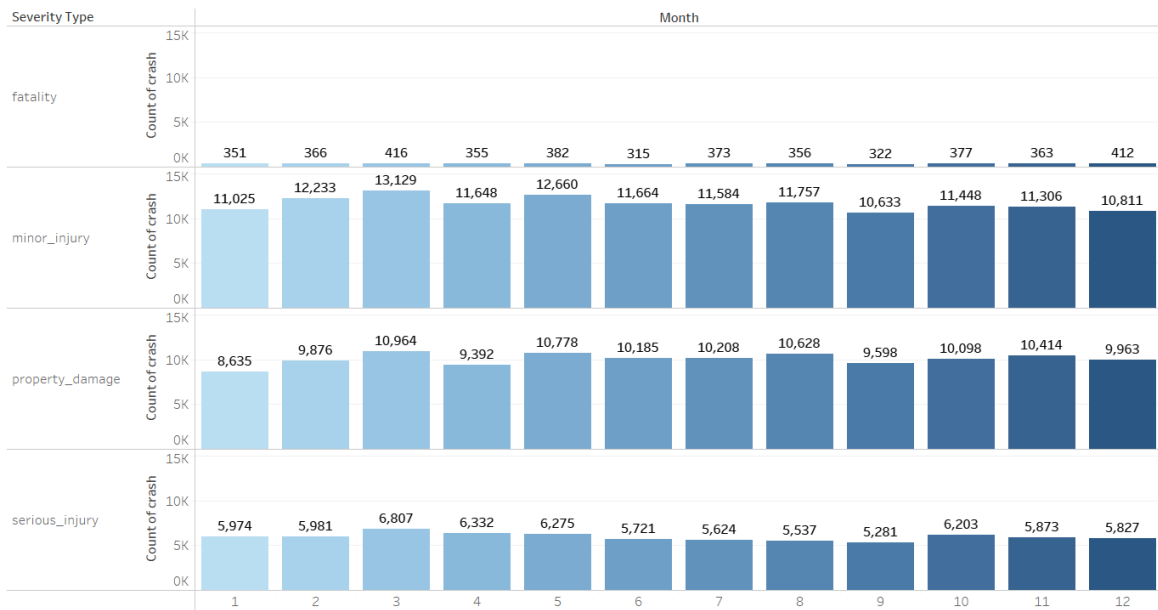
Wnioski:

Myślę że powodem takiej różnicy w liczbie bezwzględnej a odsetku jest to że na styczeń przypada wiele świąt (Nowy rok, dzień Australii) oraz wakacje w Austalii zaczynają od 20 grudnia i kończą 29 stycznia. W ten okres(20.12 – 29.01) ludzi rodzinami wyjeżdżają w różne miejsca Australii i z tego powodu mamy więcej wypadków.

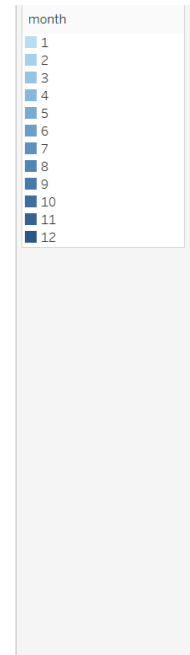
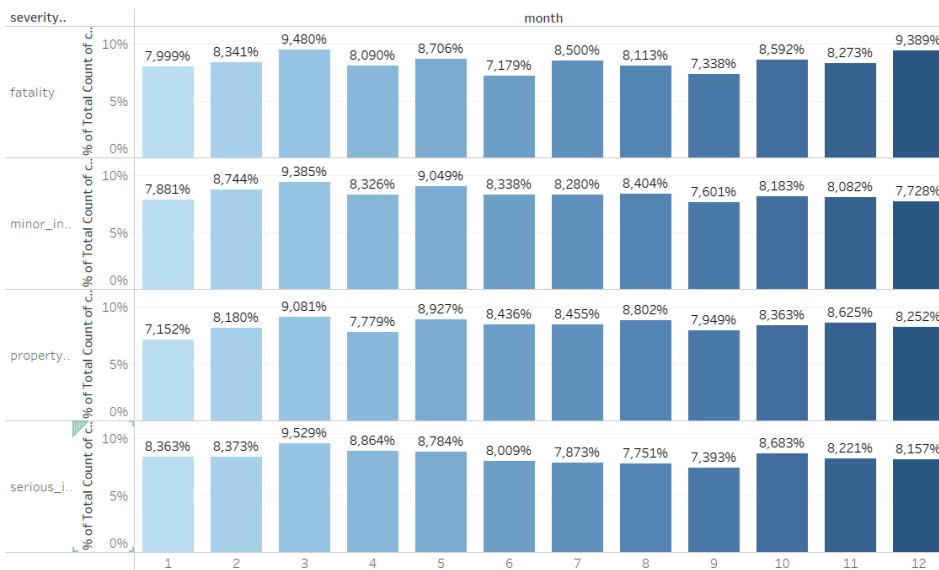
2. Ile było wypadków w zależności od miesiąca oraz surowości (“severity”) w latach 2006-2019.

Severity Type	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
fatality	351	366	416	355	382	315	373	356	322	377	363	412
minor_injury	11,025	12,233	13,129	11,648	12,660	11,664	11,584	11,757	10,633	11,448	11,306	10,811
property_damage	8,635	9,876	10,964	9,392	10,778	10,185	10,208	10,628	9,598	10,098	10,414	9,963
serious_injury	5,974	5,981	6,807	6,332	6,275	5,721	5,624	5,537	5,281	6,203	5,873	5,827

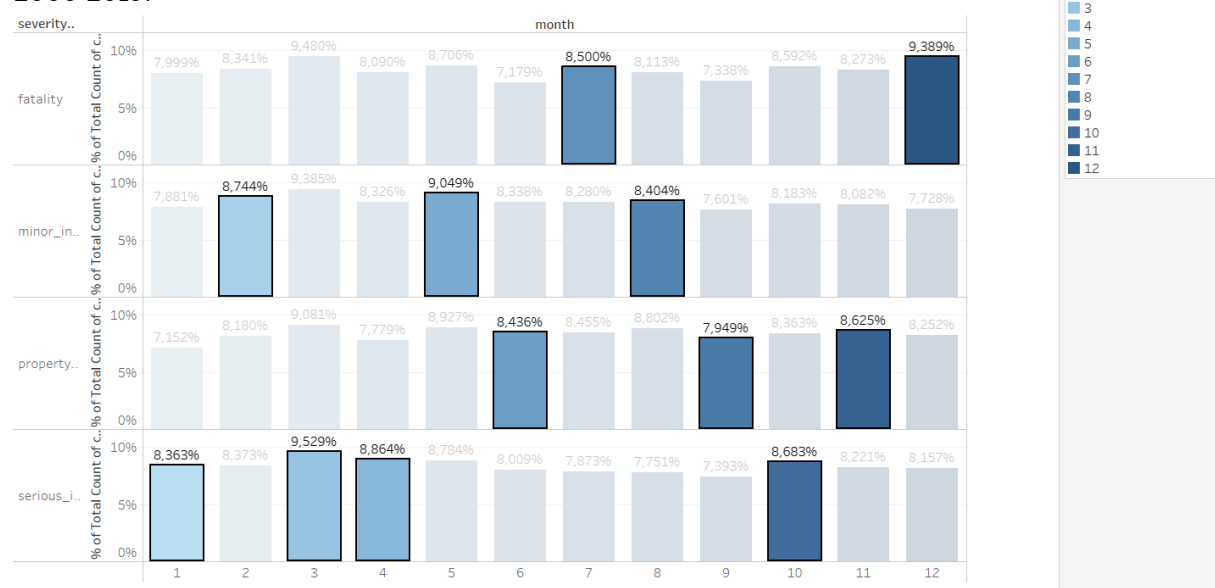
Ile było wypadków w zależności od miesiąca oraz surowości ("severity") w latach 2006-2019.



Ile było wypadków w zależności od miesiąca oraz surowości ("severity") w latach 2006-2019.



Ile było wypadków w zależności od miesiąca oraz surowości ("severity") w latach 2006-2019.



Obserwacje:

Na pierwszych dwóch wykresach możemy zobaczyć że bezwzględnie najwięcej wypadków było same z surowością „minor_injury”, dlatego skorzystaliśmy z otsetków i tutaj już jest wszystko nie tak jednoznacznie. Na ostatnim wykresie odznaczony jest maksymalny otsetek dla każdego miesiąca. Najwięcej wypadków w grudniu jest z surowością serious injury a w styczniu z surowością fatality.

Wnioski:

W grudniu i styczniu mamy wiele świąt np. Nowy Rok, czy Boże Narodzenie. Z tego powodu mamy więcej wypadków z surowością fatality oraz serious injury.

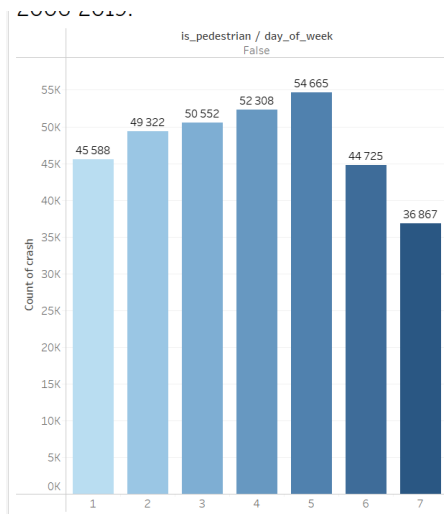
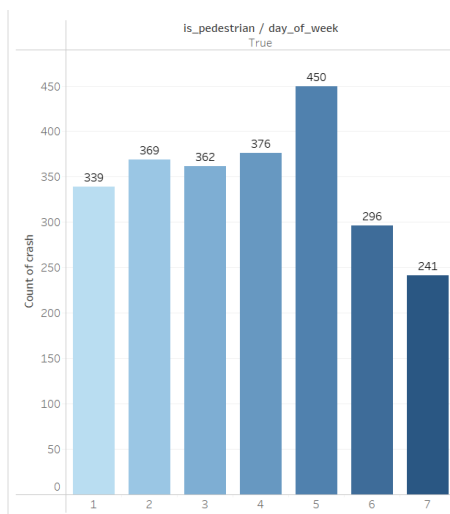
3. Ile było wypadków w zależności od dnia tygodnia oraz uczestnictwa pieszych w latach 2006-2019.

	day_of_week						
is_pedes..	1	2	3	4	5	6	7
False	45 588	49 322	50 552	52 308	54 665	44 725	36 867

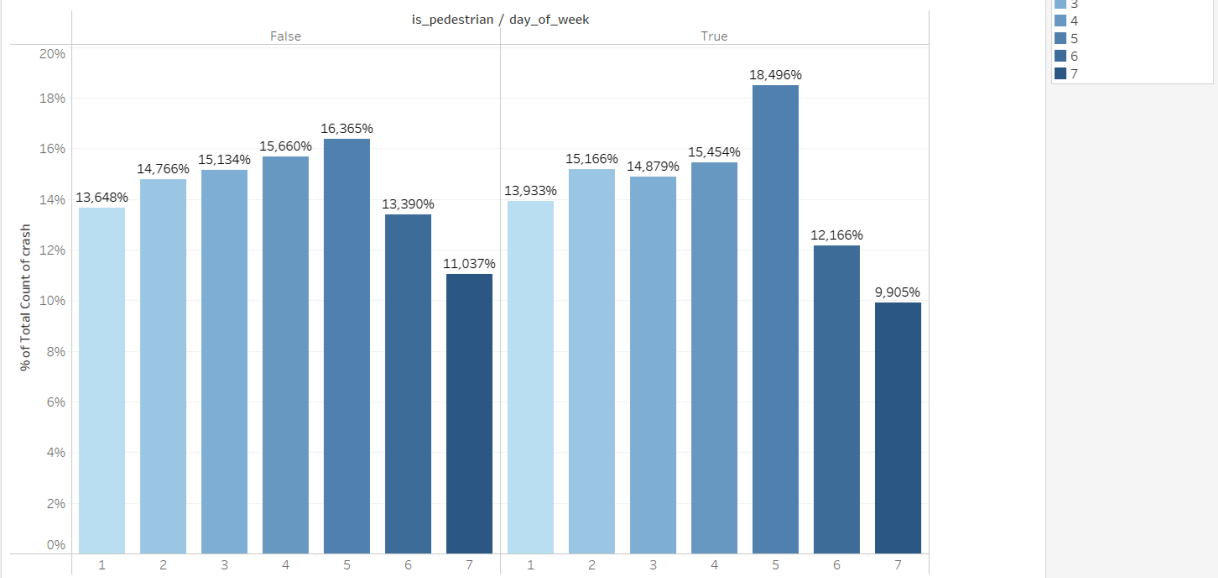
	day_of_week						
is_pedes..	1	2	3	4	5	6	7
True	339	369	362	376	450	296	241

	day_of_week						
is_pedes..	1	2	3	4	5	6	7
False	13,648%	14,766%	15,134%	15,660%	16,365%	13,390%	11,037%
True	13,933%	15,166%	14,879%	15,454%	18,496%	12,166%	9,905%

	day_of_week						
is_pedes..	1	2	3	4	5	6	7
False	13,648%	14,766%	15,134%	15,660%	16,365%	13,390%	11,037%
True	13,933%	15,166%	14,879%	15,454%	18,496%	12,166%	9,905%



Ile było wypadków w zależności od miesiąca oraz surowości ("severity") w latach 2006-2019.



Obserwacje:

Na 1 wykresie oczwicie bezwzględnie więcej wypadków bez pieszych, dlatego popatrzymy na 2 wykres na którym mamy wszystko w odsetkach. Tutaj widzimy że największy odsetek jest w piątek i co najważniejsze, że ten odsetek jest z pieszymi, a 2 najmniejszych w sobotę i niedzielę .

Wyniki:

Największy odsetek mamy same w piątek, moim zdaniem dlatego że piątek to jest koniec tygodnia pracy i wsytkie ludzi starają jak najszybciej wrócić do swoich mieszkań. Z powodu że wiele ludzi nie mają swojego samochodu dlatego wypadki z pieszymi zdarzają częściej. Sobota i niedziela to dni wolne od pracy, ludzi często siedzą w swoich mieszkaniach(dlatego same w ogóle mniej wypadków niż w inne dni tygodnia) i spacerują, czy jadą dokądś (dlatego więcej wypadków same bez uczestnictwa pieszych).

4. Ile było wypadków w zależności dnia tygodnia oraz pogody w latach 2006-2019

day_of_...	weather_type					
	fine	fog	high_wind	rain	smoke_dust	snow
1	13,56%	12,47%	12,42%	12,56%	13,66%	7,79%
2	14,50%	15,28%	16,10%	12,85%	12,33%	18,18%
3	14,77%	14,49%	14,35%	13,69%	13,22%	23,38%
4	14,96%	14,04%	15,47%	16,78%	10,43%	5,19%
5	15,52%	14,94%	14,87%	17,78%	12,92%	9,09%
6	14,10%	15,22%	13,08%	14,84%	20,56%	18,18%
7	12,60%	13,54%	13,71%	11,50%	16,89%	18,18%



Obserwacje:

W tym przypadku nie możemy wyróżnić tylko jedną anomalię związaną z pogodą snow w środek oraz czwartek.

Wnioski:

Z tego powodu że mamy dataset z Australii mogę tylko powiedzieć że w Australii śnieg pada zwykle na północy Australii a w innych regionach bardzo rzadko. Z ostatniego zdjęcia możemy zobaczyć że stany „WA – Western Australia” oraz „VIC – Victoria” znajdują na południu Australii. Według mnie to jedyne z czym może być związana taka anomalia.

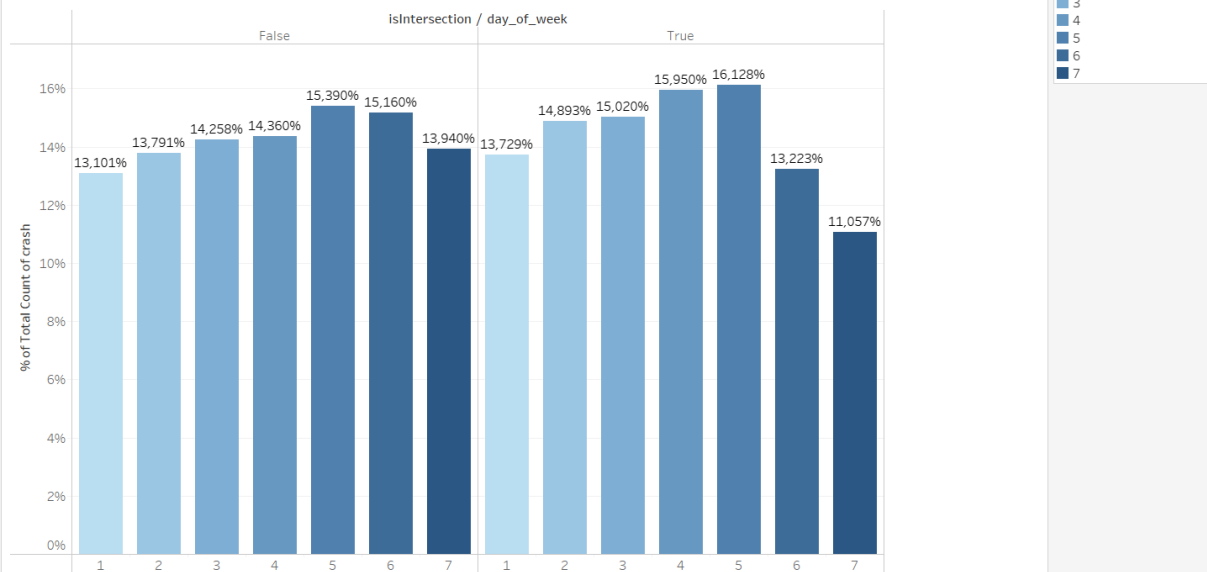
5. Ile było wypadków w zależności od stanu oraz surowości „intersection” w latach 2006-2019.

isInterse..	day_of_week						
	1	2	3	4	5	6	7
False	11 706	12 323	12 740	12 831	13 752	13 546	12 456
True	11 906	12 915	13 025	13 832	13 986	11 467	9 589

isInterse..	day_of_week						
	1	2	3	4	5	6	7
False	13,101%	13,791%	14,258%	14,360%	15,390%	15,160%	13,940%
True	13,729%	14,893%	15,020%	15,950%	16,128%	13,223%	11,057%

isInterse..	day_of_week						
	1	2	3	4	5	6	7
False	13,101%	13,791%	14,258%	14,360%	15,390%	15,160%	13,940%
True	13,729%	14,893%	15,020%	15,950%	16,128%	13,223%	11,057%

Ile było wypadków w zależności od miesiąca oraz surowości ("severity") w latach 2006-2019.



Obserwacje:

Patrząc na pierwsze 3 wykresy możemy ostatecznie powiedzieć że od poniedziałku i aż do piątku mamy więcej wypadków z intersection True ale w weekendy sytuacja zmienia i mamy już więcej wypadków same z intersection False.

Wyniki:

Dlaczego tak jest? Moim zdaniem Od poniedziałku do piątku mamy więcej wypadków same z intersecityon True ponieważ w dni robocze ludzi więcej jeżdżą w przedziału miasta, a w weekendy wyjeżdżają za miasto np. Na spacer w las czy góry.

6. Ile było wypadków w zależności od godziny oraz „intersection” w latach 2006-2019.

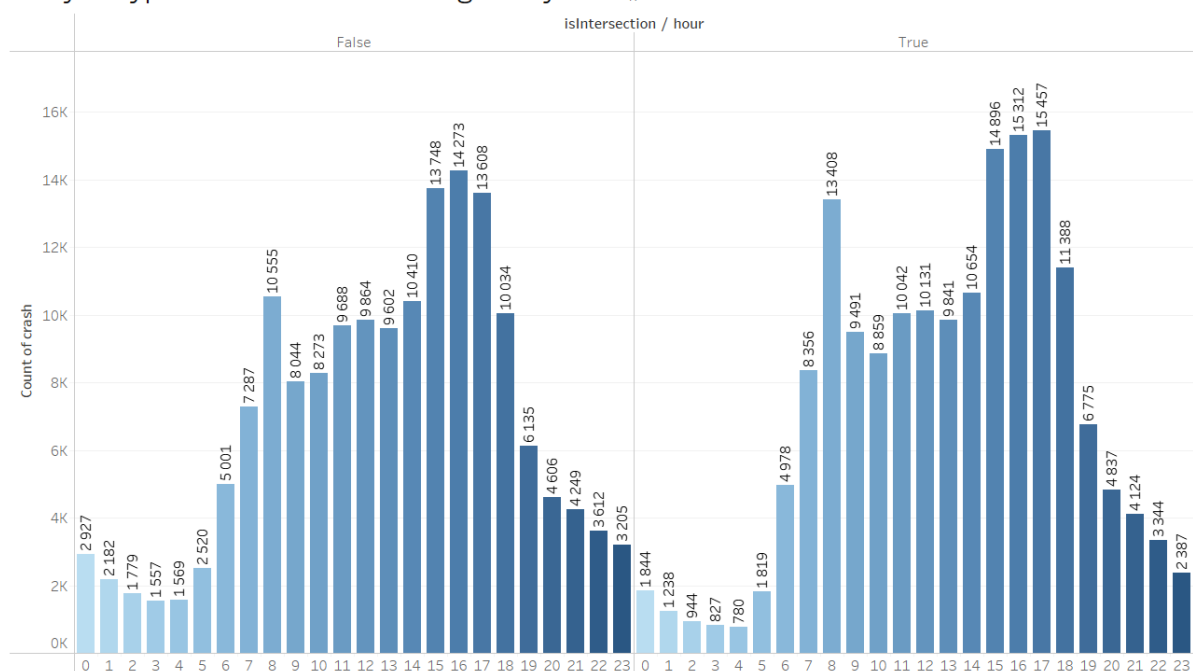
isInt..	hour																							
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
False	2 927	2 182	1 779	1 557	1 569	2 520	5 001	7 287	10 555	8 044	8 273	9 688	9 864	9 602	10 410	13 748	14 273	13 608	10 034	6 135	4 606	4 249	3 612	3 205
True	1 844	1 238	944	827	780	1 819	4 978	8 356	13 408	9 491	8 859	10 042	10 131	9 841	10 654	14 896	15 312	15 457	11 388	6 775	4 837	4 124	3 344	2 387

IsInterse..	0	1	2	3	4	5	6	7	8	hour	9	10	11	12	13	14	15	16	17	18
False	1.858%	1.609%	1.333%	1.218%	1.118%	1.593%	2.946%	4.095%	5.756%	4.771%	4.896%	5.689%	6.013%	5.934%	6.241%	7.890%	8.062%	7.827%	6.380%	
True	1.219%	1.027%	0.808%	0.714%	0.607%	1.046%	2.768%	4.446%	7.302%	5.603%	5.128%	5.572%	5.777%	5.758%	5.829%	7.973%	7.789%	8.522%	7.220%	
<div><div></div>																				

19	20	21	22	23
4.140%	2.978%	2.880%	2.498%	2.276%
4.512%	3.239%	2.889%	2.445%	1.808%

isInterse..	6	7	8	9	10	11	12	13	14	hour	15	16	17	18	19	20	21
False	2.946%	4.095%	5.756%	4.771%	4.896%	5.689%	6.013%	5.934%	6.241%	7.890%	8.062%	7.827%	6.380%	4.140%	2.978%	2.880%	
True	2.768%	4.446%	7.302%	5.603%	5.128%	5.572%	5.777%	5.758%	5.829%	7.973%	7.789%	8.522%	7.220%	4.512%	3.239%	2.889%	
<div><div></div>																	

Ile było wypadków w zależności od godziny oraz „intersection” w latach 2006-2019.



Obserwacje:

Z pierwszych dwóch wykresów widzimy że znaczenia idą prawie takie same ale różnica jest. Mamy kilka anomalnych godzin, a same 8, 15, 16, 17 oraz 18. Jednak mamy różnice między odsetkami z intersection True oraz False np. o 8 mamy aż na 1.6% więcej z intersection True, podobną różnicę możemy zobaczyć o 17 gdzie różnica to 0.7%.

Wnioski:

Te anomalie są nie czym innym jak po prostu czas pik. O ósmej ludzi jadą na pracę a od 15-18 zwracają do domu. Różnica jest wywołana tym że jeśli

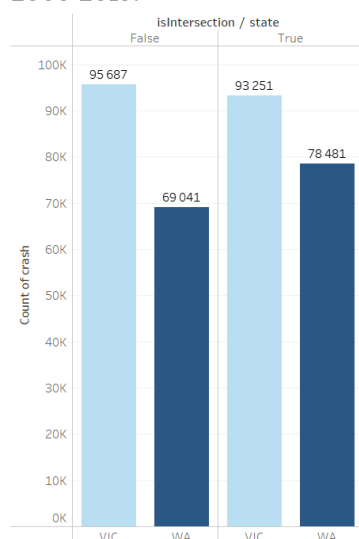
popatrzymy na wykres to zobaczymy że aż do 7:00 mamy większe odsetki same z intersection False, według mnie wyzwane są tym że ludzi poza miastem jadą do miasta i wtedy więcej wypadków zdarzają z intersection False. O 7:00 ludzi już po trochu przejeżdżają do miasta i wypadków z intersection True już zdarzają więcej, ale pik wypadków rano mamy o 8:00 kiedy już na drogach możemy zobaczyć tory.

7. Ile było wypadków w zależności od stanu ("state") oraz „intersection” w latach 2006-2019.

isInterse..	state	
	VIC	WA
False	95 687	69 041
True	93 251	78 481

isInterse..	state	
	VIC	WA
False	50,645%	46,800%
True	49,355%	53,200%

Ile było wypadków w zależności od stanu ("state") oraz „intersection” w latach 2006-2019.



Obserwacje:

Na początku możemy zobaczyć że więcej wypadków między stanami z intersection True zdarza się same w stanie VIC, ale jeśli popatrzymy w odsetkach to możemy zrozumieć że odsetkowo w stanie WA z intersection True zdarzają więcej wypadków niż w stanie VIC z intersection True.

Wnioski:

Nie ma zależności

8. Ile było wypadków w zależności uczestnictwa pieszych oraz surowości ("severity") w latach 2006-2019.

is_pedestrian	severity_type			
	fatality	minor_injury	property_damage	serious_injury
False	4 309	139 278	119 655	70 785
True	79	620	1 084	650

is_pedes..	severity_type			
	fatality	minor_injury	property_da..	serious_inju..
False	98,20%	99,56%	99,10%	99,09%
True	1,80%	0,44%	0,90%	0,91%

Obserwacje:

Jeśli popatrzymy na pierwszy wykres możemy zobaczyć że we wszystkich wypadkach mamy więcej wypadków z is_pedestrian False, ale to są liczbe bezwzględne i musimy popatrzyć na ten wykres odsetkowo. Na drugim wykresie już widocznie że odsetkowo mamy więcej fatality, serious_injury same kiedy uczestniczyli piesze.

Wyniki:

Moim zdaniem mamy mniej wypadków z surowością minor_injury, po prostu dlatego że kiedy w wypadku uczestniczy pieszy to mniej jest szansa że człowiek będzie w dobrym stanie po wypadku. **Wyniki pokazują, że średnie ryzyko poważnych obrażeń u pieszego uderzonego przez pojazd sięga 10% przy prędkości 16 mil na godzinę, 25% przy 23 mil na godzinę, 50% przy 31 mil na godzinę, 75% przy 39 mil na godzinę i 90% przy 46 km/h mph. Średnie ryzyko śmierci pieszego sięga 10% przy prędkości uderzenia 23 mil na godzinę, 25% przy 32 milach na godzinę, 50% przy 42 milach na godzinę, 75% przy 50 milach na godzinę i 90% przy 58 milach na godzinę.**

Jeszcze mamy bardzo interesujący wynik że `property_damage` = 0.9%. Według mnie to mogą być wypadki kiedy kierowca zobaczył pieszego i spróbował skierować w inną stronę i np. zjechał z drogi.

Informacja odznaczona na grubo wzięta z <https://aaafoundation.org/impact-speed-pedestrians-risk-severe-injury-death/>

9. Ile było wypadków w zależności surowości(“severity”) oraz „intersection” w latach 2006-2019.

isInterse..	severity_type			
	fatality	minor_injury	property_damage	serious_injury
False	3 195	64 973	56 949	39 611
True	1 193	74 925	63 790	31 824

isInterse..	severity_type			
	fatality	minor_injury	property_damage	serious_injury
False	72,81%	46,44%	47,17%	55,45%
True	27,19%	53,56%	52,83%	44,55%

Obserwacje:

Na pierwszym wykresie nie widzimy nic anomального dlatego popatrzymy na drugi i zobaczymy dane w odsetkach. Z odsetków bardzo dobrze widoczna anomalia z surowością „fatality” oraz bez skrzyżowania, tych wypadków jest na 45% więcej niż wypadków na skrzyżowaniu!

Wnioski:

Według mnie więcej samych fatalnych wypadków bez skrzyżowania może zdarzać z bardzo wielu powodów np.:

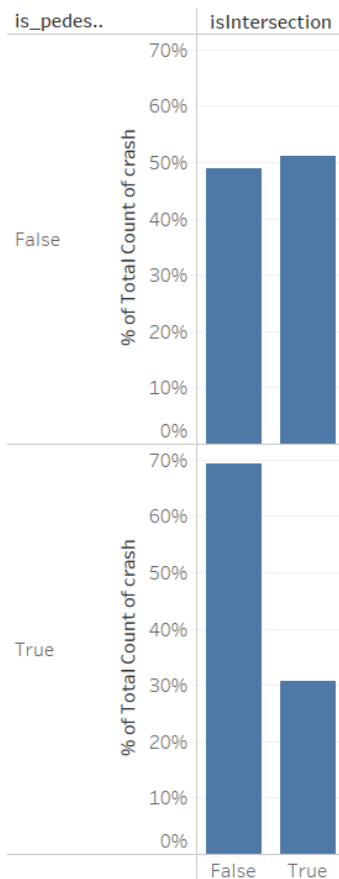
1. Na skrzyżowaniu kierowcy jeżdżą z bardzo mniejszą prędkością, niż na drodze bez skrzyżowania.
2. Na skrzyżowaniu kierowcy więcej uważają niż bez

Jest jeszcze wiele powodów ale nie możemy stwierdzać na pewno dlatego że nie wiemy czy te wypadki bez skrzyżowania byli w mieście czy poza miastem, czy kierowca wyjechał na część pieszą czy nie, może pieszy przechodził drogę w nieprawidłowym miejscu. Możemy tylko zasugerować że było to z tych czy innych powodów ale nie stwierdzać na pewno

10. Ile było wypadków w zależności od uczestnictwa pieszych oraz „intersection” w latach 2006-2019.

isIntersection	is_pedestrian	
	False	True
False	163 041	1 687
True	170 986	746

isInterse..	is_pedestrian	
	False	True
False	48,81%	69,34%
True	51,19%	30,66%



Obserwacje:

Bezwzględnie oczywiście mamy więcej wypadków bez uczestnictwa pieszych, dlatego zobaczymy w odsetkach i popatrzymy na drugi wykres. Możemy tutaj zobaczyć interesującą rzecz, wypadków z uczestnictwem pieszych więcej same z intersection False, a wypadków bez pieszych więcej z intesection True.

Wyniki:

Moim zdaniem więcej wypadków z pieszymi odbywają same z intersection False z dwóch powodów:

1. Dlatego że kierowcy więcej uważają same na skrzyżowaniu dróg.
2. Dlatego że na skrzyżowaniu dróg, sama droga jest bardzo większa niż na zwykłej drodze gdzie mamy 2-4 rzędy i ludzie nie chcą i przechodzić drogę na czerwone światelko. Również jakiś ludzi mogą po prostu przebiegać drogę na odcinku bez sygnalizacji świetlnej.

Wnioski:

Przeprowadziliśmy analizy naszej hurtowni danych związanej z wypadkami na drogach Australii. Wykorzystaliśmy model danych gwiazda ze względu na szbszyj dostęp do danych. Skorzystaliśmy z narzędzia do wizualizacji danych „Tableau”. W niektórytch przypadkach byliśmy zmuszeni nie patrzeć na liczby bezwzględne, dlatego że liczby bezwzględne nie zawsze dają prawidłowy obraz do analizy, tylko obliczyć odsetkowe wartości i już korzystając z nich zrobić jakieś obserwacje oraz wnioski.

Bibliografia:

Wszystkie dane są wzięte z [kaggle](#).

Informacja o prędkości: <https://aaaafoundation.org/impact-speed-pedestrians-risk-severe-injury-death/>

Mapa Australii ze stanami: <http://www.beforeiforget.co.uk/2011/north/>