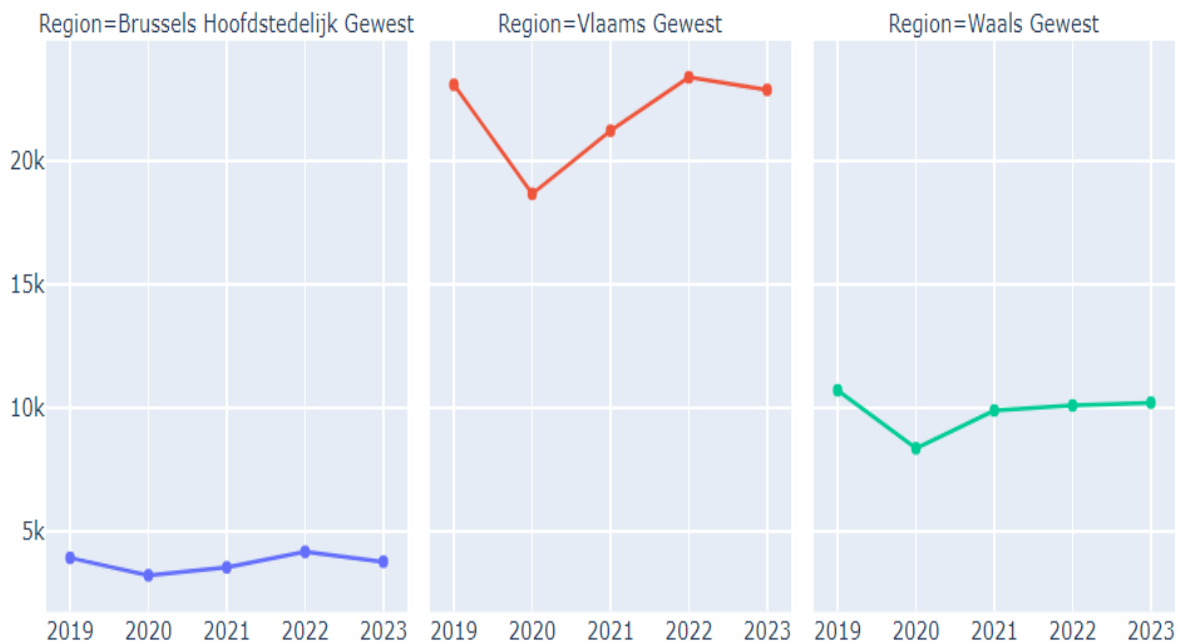




Verkeersongevalsanalyse in België (2019-2023)

Een integrale data-analyse van verkeersveiligheid en ongevallen



Gokarna Bhusal
Campus Syntra AB, Leuven

01.04.2025



Inhoudsopgave

Beschrijving.....	2
1. Inleiding.....	3
1.1. Doel van het project	3
1.1. Achtergrond en relevantie.....	3
2. Data Verzameling en Verwerking.....	4
2.1. Datainzameling via de Belgische open data (Statbel).....	4
2.2. Implementatie van de data-inleespipeline	4
2.3. Data cleaning, validatie en transformatie.....	4
3. Data Analyse.....	12
3.1. Urenanalyse	12
3.2. Dagelijkse Ongevalanalyse:	12
3.3. Maandelijkse Ongevallenanalyse.....	13
3.4. Jaarlijkse trends in ongevallen	14
4. Regional Ongeval analyse:	15
4.1. Gewestelijk analyse:	15
4.2. Provinciaal analyse:	16
4.3. Gemeentelijke analyses.....	16
4.4. Analyse van ongevallen in bebouwde versus niet-bebouwde gebieden.....	17
5. Wegtype Analyse	18
5.1. Snelwegen versus hoofdwegen	18
5.2. Overzicht van ongevallen, ernstige verwondingen en fataliteiten	19
5.3. Risico's en verkeersintensiteit op verschillende wegtypen	19
6. Analyse van Lichtcondities	20
7. Analyse van Botsingstypen	21
7.1. Overzicht en trends per botsingstype:	21
8. Zakelijke vragen en mogelijke antwoorden:	22
8.1. Welke factoren hebben de grootste invloed op het aantal en de ernst van verkeersongevallen in België?.....	22
8.2. Hoe verschillen verkeersongevallen per regio en tijdstip, en zijn er bepaalde trends te ontdekken?	23
9. Samenvatting en Conclusies	24

Beschrijving

Dit project richt zich op de analyse van verkeersdata in België over de periode 2019 tot 2023. De studie omvat:

- Het inladen, opschonen en valideren van data afkomstig uit verschillende bronnen en formaten.
- Een grondige data-analyse met focus op tijdsgebonden patronen (uur-, dag-, maand- en jaarverdelingen), regionale verschillen en uiteenlopende wegtypes.
- Een specifieke analyse van lichtcondities, botsingstypen en de ernst van ongevallen.
- Vergelijkingen tussen snelwegen, hoofdwegen en andere wegen, met aandacht voor verkeersintensiteit, letsel- en sterftcijfers.
- Een evaluatie van de impact van seizoenen, dagelijkse routines en specifieke locaties (bijvoorbeeld in stads- versus landelijke gebieden).

Op basis van de verkregen inzichten worden daarnaast gerichte aanbevelingen gedaan voor zowel stedelijke als landelijke verkeersveiligheidsmaatregelen. Het doel van deze analyse is bij te dragen aan een beter begrip van de risicofactoren en om de basis te leggen voor maatregelen die de verkeersveiligheid in België verder kunnen verbeteren.

1. Inleiding

1.1. Doel van het project

Het doel van dit project is om een diepgaande analyse uit te voeren van verkeersongevallen in België over de periode 2019 tot en met 2023. Door gebruik te maken van open data van betrouwbare bronnen zoals Statbel, tracht dit project de trends, patronen en risicofactoren die bijdragen aan verkeersincidenten te identificeren. Hierbij worden zowel temporele (uur-, dag-, maand- en jaardistributie) als ruimtelijke variaties (regionale verschillen en stedelijke versus landelijke omstandigheden) geanalyseerd.

De reikwijdte van het project omvat verschillende stappen:

- **Data-inzameling en integratie:** Het inlezen van data uit diverse bestandsformaten (zoals CSV, XLSX, TXT en JSON) en het samenvoegen van deze bronnen tot één consistente dataset.
- **Data cleaning en validatie:** Het opschonen en transformeren van de data, waaronder het hernoemen van kolommen, het verwijderen van overbodige informatie, het corrigeren van datatypeproblemen en het uitvoeren van feature engineering op bijvoorbeeld datumvelden.
- **Gedetailleerde data-analyse:** Het onderzoeken van ongevallen op basis van tijd (uur, dag, maand, jaar), wegtypen, lichtcondities, botsingstypen, en regionale verschillen.
- **Aanbevelingen:** Het formuleren van concrete aanbevelingen voor verkeersveiligheidsmaatregelen, gebaseerd op de inzichten uit de data.

Op deze manier beoogt het project niet alleen het verschaffen van een helder beeld van de verkeersveiligheidssituatie in België, maar ook het ondersteunen van besluitvorming bij overheden en andere stakeholders om de wegen veiliger te maken.

1.1. Achtergrond en relevantie

Verkeersongevallen vormen wereldwijd een belangrijke maatschappelijke uitdaging, waarbij de gevolgen variëren van materiële schade tot ernstig letsel en zelfs verlies van mensenlevens. In België, met zijn complexe wegennetwerk en hoge verkeersintensiteit, is het analyseren van deze incidenten van cruciaal belang. De afgelopen jaren hebben ertoe geleid dat verkeersveiligheid een prioriteit is geworden voor zowel beleidsmakers als de samenleving.

De achtergrond voor dit onderzoek ligt in de noodzaak om op basis van data-inzichten de onderliggende oorzaken van ongevallen beter te begrijpen. Factoren zoals veranderende verkeersvolumes, diverse wegtypen en seizoensgebonden invloeden dragen bij aan de complexiteit van verkeersincidenten. Bovendien hebben externe invloeden, zoals de COVID-19-pandemie in 2020, geleid tot opmerkelijke veranderingen in verkeersstromen en ongevallenpatronen.

De relevantie van dit project is tweeledig. Enerzijds biedt het waardevolle inzichten die kunnen bijdragen aan een effectievere inzet van verkeersveiligheidsmaatregelen en infrastructuurverbeteringen. Anderzijds kan de verkregen kennis beleidsmakers en hulpdiensten ondersteunen bij het ontwikkelen van preventieve strategieën, zoals gerichte voorlichtingscampagnes en verbeteringen in stedelijk en landelijk verkeersbeheer. Door de risico's en oorzaken van verkeersongevallen te analyseren, helpt dit onderzoek bij het streven naar een veiligere verkeersomgeving, wat uiteindelijk leidt tot minder slachtoffers en een hogere verkeersveiligheid voor iedereen.

2. Data Verzameling en Verwerking

2.1. Datainzameling via de Belgische open data (Statbel)

Ik heb de Belgische open source website (<https://statbel.fgov.be/>) gebruikt om gegevens te downloaden. Ik wilde niet alleen gedurende 1 jaar, maar de laatste 5 jaar analyses uitvoeren om de trend van verkeersongevallen in België per jaar te zien. Als de laatste gegevens die ik in 2023 vind, gebruik ik gegevens van 2019 tot 2023 om dit project te doen.



De gegevens zijn opgeslagen in verschillende formaten (.xlsx en .txt). De eerste stap was om deze verschillende formaten en bestanden te lezen en ze samen te voegen in één CSV-bestand, zodat het inleesproces van de data eenvoudiger wordt.

2.2. Implementatie van de data-inleespipeline

Interfaces voor Data Inladen en Schoonmaken

Om de leesbaarheid en onderhoudbaarheid van de code te verbeteren, maak ik gebruik van het *Separation of Concerns (SoC)*-principe. Dit houdt in dat we verschillende verantwoordelijkheden in aparte modules of interfaces onderbrengen.

Interface voor Data Inladen

De interface **LoadDataProcess** definieert de methoden die gerelateerd zijn aan data-invoer. Dit zorgt ervoor dat alle mechanismen voor het lezen van data uit diverse bronnen (bestanden, databases, API's, etc.) voldoen aan een consistente contract.

Python code:

```
class LoadDataProcess(ABC):  
    @abstractmethod  
    def read_data(self):  
        pass
```

2.3. Data cleaning, validatie en transformatie

Interface voor Data validatie

De interface **DataCleaningProcess** specificeert de methoden voor het valideren en opschonen van de data. De focus ligt hierbij op het consistent uitvoeren van data preprocessing, zoals validatie, transformatie en feature engineering.

Python code:

```

class DataCleaningProcess(ABC):
    @abstractmethod
    def validate_data(self):
        pass

    @abstractmethod
    def rename_columns(self):
        pass

    @abstractmethod
    def transform_data(self):
        pass

    @abstractmethod
    def print_information(self):
        pass

```

Uitleg van de Interfaces

LoadDataProcess Interface:

Verantwoordelijkheid: Definieert methoden voor data-invoer.

Doel: Zorgt voor het lezen van data van diverse bronnen (bijv. bestanden, databases, API's) en garandeert dat alle mechanismen voor het inlezen van de data werken volgens een vast contract.

DataCleaningProcess Interface:

Verantwoordelijkheid: Specificeert methoden voor data validatie en opschoning.

Doel: Beheert taken op het gebied van data preprocessing (zoals validatie, transformatie en feature engineering) en zorgt zo voor consistentie in de data-cleansing processen.

Door deze verantwoordelijkheden te scheiden, zorgen we ervoor dat een wijziging in één gebied (bijvoorbeeld het toevoegen van een nieuwe databron) niet onbedoeld een ander gebied (zoals de logica voor data opschoning) beïnvloedt.

Pipeline

Voor deze stap maak ik een inleespipeline. Dit maakt mijn code zeer flexibel, zodat verschillende soorten formaten en bronnen eenvoudig kunnen worden ingelezen.

Python code:

```
class FileHandler:
```

```
    def __init__(self, path):
        self.path = path

        # Roep de read_data() methode aan om alle data in te laden en samen te voegen.
        self.data = self.read_data()
```

```
    def read_data(self):
```

```
        data_list = []
```

```
        # Controleer of het opgegeven pad een directory (map) is.
```

```
        if os.path.isdir(self.path):
```

```
            for file in os.listdir(self.path):
```

```
                file_path = os.path.join(self.path, file)
```

```
                # Lees de data op basis van de bestandsextensie
```

```
                if file.endswith(".txt"):
```

```
                    # Voor .txt-bestanden gaan we ervan uit dat het scheidingsteken '|' is.
```

```
                    data_list.append(pd.read_csv(file_path, sep='|', dtype=str))
```

```
                elif file.endswith(".csv"):
```

```
                    data_list.append(pd.read_csv(file_path, dtype=str))
```

```
                elif file.endswith(".xlsx"):
```

```
                    data_list.append(pd.read_excel(file_path, dtype=str))
```

```
                elif file.endswith(".json"):
```

```
                    data_list.append(pd.read_json(file_path))
```

```
                else:
```

```
                    print(f"Bestandsformaat niet ondersteund: {file}")
```

```
        else:
```

```
            raise ValueError("Het opgegeven pad is geen map.")
```

```
        # Voeg alle ingelezen data samen in één DataFrame en reset de indices
```

```
        data = pd.concat(data_list, ignore_index=True)
```

```
return data
```

Kolommen van de Samengevoegde DataFrame

Na het samenvoegen bevat je DataFrame de volgende kolommen:

```
['DT_DAY', 'DT_HOUR', 'CD_DAY_OF_WEEK', 'TX_DAY_OF_WEEK_DESCR_FR',  
'TX_DAY_OF_WEEK_DESCR_NL', 'CD_BUILD_UP_AREA',  
'TX_BUILD_UP_AREA_DESCR_NL', 'TX_BUILD_UP_AREA_DESCR_FR',  
'CD_COLL_TYPE', 'TX_COLL_TYPE_DESCR_NL', 'TX_COLL_TYPE_DESCR_FR',  
'CD_LIGHT_COND', 'TX_LIGHT_COND_DESCR_NL', 'TX_LIGHT_COND_DESCR_FR',  
'CD_ROAD_TYPE', 'TX_ROAD_TYPE_DESCR_NL', 'TX_ROAD_TYPE_DESCR_FR',  
'CD_MUNTY_REFNIS', 'TX_MUNTY_DESCR_NL', 'TX_MUNTY_DESCR_FR',  
'CD_DSTR_REFNIS', 'TX_ADM_DSTR_DESCR_NL', 'TX_ADM_DSTR_DESCR_FR',  
'CD_PROV_REFNIS', 'TX_PROV_DESCR_NL', 'TX_PROV_DESCR_FR',  
'CD_RGN_REFNIS', 'TX_RGN_DESCR_NL', 'TX_RGN_DESCR_FR', 'MS_ACCT',  
'MS_ACCT_WITH_DEAD', 'MS_ACCT_WITH_DEAD_30_DAYS',  
'MS_ACCT_WITH_MORY_INJ', 'MS_ACCT_WITH_SERLY_INJ',  
'MS_ACCT_WITH_SLY_INJ']
```

De original dataset bevat **35 kolommen** waarin voor dezelfde informatie kolommen aanwezig zijn in drie talen: Engels, Nederlands en Frans. Dit betekent dat voor elk gegeven er drie kolommen bestaan die dezelfde informatie bevatten, maar in verschillende talen.

Doelstelling:

- Alleen één set kolommen behouden (bijvoorbeeld de Nederlandse versie).
- Deze kolommen hernoemen naar meer leesbare en begrijpelijke namen.

Daarnaast bevat de dataset:

- 176.277 rijen.
- De datatypes van de meeste kolommen zijn object.
- Er ontbreken waarden in enkele belangrijke kolommen zoals de bouwomgeving (Build-up area), het soort botsing (coll_type), lichtconditie (Light Condition), wegtype (Road-Type) en provincie.

Daarom is het noodzakelijk om een grondige cleaning, validatie en transformatie van de data uit te voeren. Deze stappen worden gecoördineerd in het bestand `src/verkeerOngevallen/models/data_processor.py`. Het ontwerp bevordert bovendien.

Polymorfisme, zodat implementaties eenvoudig onderling verwisselbaar zijn zonder de rest van het systeem te beïnvloeden.

Stappenplan voor Data Cleaning en Validatie

1. **Kiezen van de Taal:** Omdat we drie versies van dezelfde informatie hebben, kiezen we één taal (bijvoorbeeld Nederlands) om duplicaties te voorkomen. *Waarom?* Dit vereenvoudigt de data-analyse en verbetert de consistentie binnen de dataset.
2. **Hernoemen van Kolommen:** De oorspronkelijke kolomnamen kunnen technisch en onduidelijk zijn. Door ze te hernoemen naar simpelere en intuïtieve namen, wordt het analyseren en communiceren eenvoudiger. *Waarom?* Duidelijke kolomnamen verminderen de kans op fouten bij latere analyses en verbeteren de leesbaarheid voor alle betrokkenen.
3. **Omgaan met Missende Waarden:** Er zijn missende waarden in onder andere de kolommen voor de bouwomgeving, botsingstype, lichtconditie, wegtype en provincie. Mogelijke strategieën zijn:
 - Vullen met een standaardwaarde (bijv. "Onbekend").
 - Het verwijderen van rijen bij overmatige missende data.
 - Geavanceerdere imputatietechnieken toepassen. *Waarom?* Het correct afhandelen van missende waarden voorkomt verstoringen in analyses en modellen.
4. **Transformatie van Datatypes:** Veel kolommen staan momenteel als object, wat niet ideaal is voor numerieke bewerkingen. Conversie naar de juiste datatype (bijvoorbeeld naar numeriek of datum/tijd) is van belang. *Waarom?* Correcte datatypes zorgen ervoor dat berekeningen, vergelijkingen en statistische analyses nauwkeurig en robuust worden uitgevoerd.
5. **Implementatie via Polymorfisme:** Door de logica voor data cleaning en validatie in speciale modules of klassen onder te brengen, en door interfaces te definiëren, maak gebruik van polymorfisme. *Waarom?* Dit maakt de implementatie flexibel en onderhoudbaar. Je kunt eenvoudig een nieuwe implementatie aansluiten als de eisen veranderen, zonder dat de rest van de code aangepast hoeft te worden.

Zie je beneden alle stappen in de Python code:

```
class DataProcessor():  
    def __init__(self, data):  
        # Maak een kopie van de originele DataFrame  
        self.data = data.copy()  
        # Lijst voor het bijhouden van kolommen die behouden moeten worden  
        self.keep_cols = []  
        # Kolommen die numeriek gemaakt moeten worden  
        self.numeric_cols = ['DT_HOUR', 'MS_ACCT', 'MS_ACCT_WITH_DEAD',  
                              'MS_ACCT_WITH_DEAD_30_DAYS',  
                              'MS_ACCT_WITH_MORY_INJ', 'MS_ACCT_WITH_SERLY_INJ',  
                              'MS_ACCT_WITH_SLY_INJ']  
        # Kolom met datuminformatie
```

```

self.date_col = 'DT_DAY'

def validate_data(self):
    # Verwijder kolommen die niet bruikbaar zijn: kolommen die eindigen op "_FR" of
    # beginnen met "CD"
    for col in self.data.columns:
        if col.endswith("_FR") or col.startswith("CD"):
            self.data.drop(col, axis=1, inplace=True)
        else:
            self.keep_cols.append(col)
    # Houd alleen de bruikbare kolommen over
    self.data = self.data[self.keep_cols]

    # Verwijder spaties aan het begin en einde van de datumkolom
    self.data[self.date_col] = self.data[self.date_col].str.strip()

    # Feature engineering op de datumkolom: maak aparte kolommen voor jaar, maand en
    # dag van de maand
    self.data['Year'] = self.data[self.date_col].str[:4].astype('Int64', errors='ignore')
    self.data['Month'] = self.data[self.date_col].str[5:7].astype('Int64', errors='ignore')
    self.data['Day_of_Month'] = self.data[self.date_col].str[8:10].astype('Int64',
    errors='ignore')

    # Verwijder de originele datumkolom, omdat deze opgesplitst is in aparte componenten
    self.data.drop(self.date_col, axis=1

```

python code:

```

class AccidentDataPipeline:
    def __init__(self, path):
        self.path = path

    def process_data(self):
        # Maak een instantie van FileHandler om data in te lezen vanaf het opgegeven pad

```

```

reader = FileHandler(self.path)
original_df = reader.read_data()

# Maak een instantie van DataProcessor met de ingelezen originele data
processor = DataProcessor(original_df)

# Creëer een nieuwe map "transform_data" om de data op te slaan als deze nog niet
bestaat
if not os.path.exists("transform_data"):
    os.makedirs("transform_data")

# Sla eerst de originele data op op de opgegeven locatie
processor.transform_data(original_df,
"transform_data/original_Data_From_2019_2023.csv")

# Voer de data validatie en opschoning uit
clean_df = processor.validate_data()

# Sla de opgeschoonde data op op de opgegeven locatie
processor.transform_data(clean_df,
"transform_data/validated_Data_From_2019_2023.csv")

# Geef de opgeschoonde data terug
return processor.data

```

Uitleg van het Ontwerp

- **Implementatie van de DataCleaningProcess Interface:** De klasse DataProcessor is verantwoordelijk voor het schoonmaken, valideren en transformeren van de data. Doordat deze taak gescheiden is van het inlezen van de data (via FileHandler), blijft de code beter beheersbaar. *Voordeel:* Als je later veranderingen moet doorvoeren in de opschoning, dan kun je dat doen zonder afbreuk te doen aan de data-inleescomponent.
- **Compositie:** In de klasse AccidentDataPipeline gebruik je compositie door instanties van zowel FileHandler als DataProcessor te bezitten. Hierdoor:
 - **Verblijft de Verantwoordelijkheid in Gespecialiseerde Klassen:** De FileHandler leest de data, en de DataProcessor voert de opschoning- en transformatietaken uit.

- **Delegatie van Taken:** Specifieke verantwoordelijkheden worden aan hun eigen klassen gedelegeerd. Dit houdt de code georganiseerd, duidelijk en makkelijk onderhoudbaar.
- **Flexibiliteit en Modulariteit:** Je kunt eenvoudig nieuwe componenten toevoegen, zoals een extra type data-inlezer (bijvoorbeeld een DatabaseReader of APIReader) of verschillende data processors (zoals AccidentDataProcessor versus TrafficDataProcessor). Hierdoor kun je de implementatie uitbreiden of aanpassen zonder de kernlogica van de pijplijn aan te passen.
- **Polymorfisme:** Het ontwerp laat ruimte voor het wisselen van implementaties. Stel je voor dat je later een andere strategie voor data cleaning wil toepassen, dan kun je gemakkelijk een andere klasse maken die de DataCleaningProcess interface implementeert en deze in de pijplijn opnemen. Dit bevordert de uitwisselbaarheid van componenten.
- **Overige Belangrijke Punten:**
 1. **Data Opslaan voor Logging en Audit:** In AccidentDataPipeline worden zowel de originele als de opgeschoonde data opgeslagen in aparte CSV-bestanden. Dit is handig voor latere analyses of debugging.
 2. **Schaalbaarheid en Onderhoudbaarheid:** Door de verantwoordelijkheden helder te scheiden met behulp van compositie, houd je de code basis overzichtelijk en wordt het aanpassen van één component eenvoudiger zonder andere modules te beïnvloeden.

Na het opschonen van onze data is het tijd om de opgeschoonde DataFrame te bekijken. De uiteindelijke DataFrame bevat **19 kolommen**:

```
[ 'Hour', 'DayOfWeek', 'BuiltUpArea', 'CollisionType', 'LightCondition',
  'RoadType', 'Municipality', 'District', 'Province', 'Region',
  'Accident', 'AccidentsWithFatalities', 'AccidentsWithFatalities30Days',
  'AccidentsWithMinorInjuries', 'AccidentsWithSeriousInjuries',
  'AccidentsWithSlightInjuries', 'Year', 'Month', 'Day_of_Month' ]
```

En er zijn **176.274** rijen in de dataset.

Zoals je ziet, hebben we nauwelijks informatie verloren. We beschikken nu over schone, gevalideerde data.

Uitgevoerde Stappen tijdens het Opschoningsproces

Tijdens dit proces heb ik de volgende stappen doorlopen:

1. **Hernoemen van de kolommen:** De oorspronkelijke kolomnamen zijn aangepast naar meer leesbare en begrijpelijke namen.

2. **Verwijderen van overbodige kolommen:** Onnodige kolommen werden verwijderd, zodat alleen de relevante informatie behouden bleef.
3. **Hercontrole van de datatypes:** Veel kolommen hadden eerder nog het datatype object. Deze zijn geconverteerd naar de juiste datatypes (bijvoorbeeld numeriek of geschikt voor datum/tijd) voor een correcte verwerking.
4. **Verwijderen van overtollige spaties:** Elke record is gecontroleerd op eventuele voor- of achtervoegsels (leading/trailing spaces) en deze zijn verwijderd.
5. **Feature engineering op de datumkolom:** De datumkolom is opgesplitst in aparte componenten, zoals jaar, maand en dag van de maand, zodat je op meerdere niveaus de data kunt analyseren.

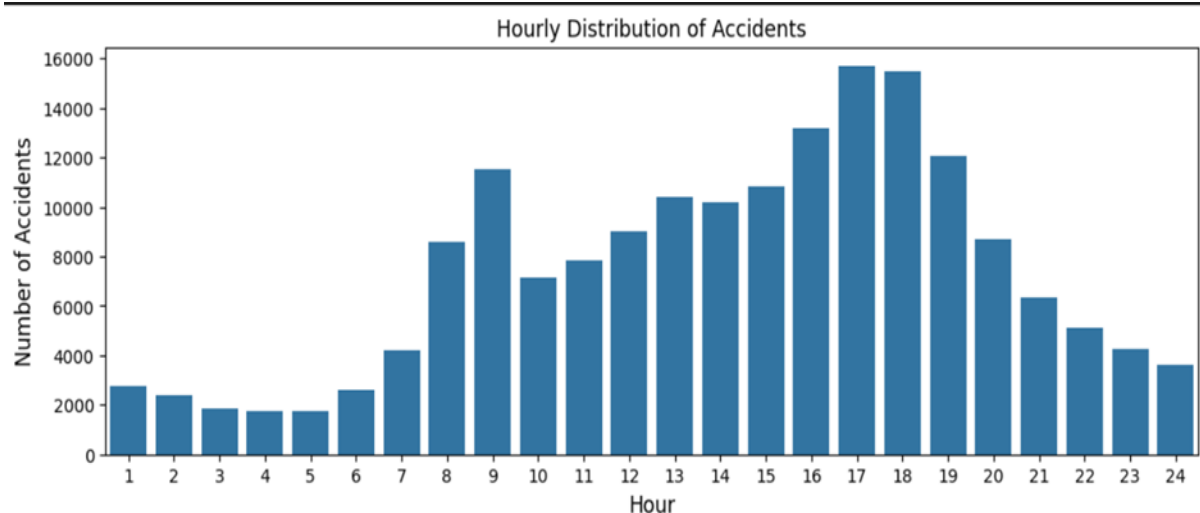
3. Data Analyse

3.1. Urenanalyse

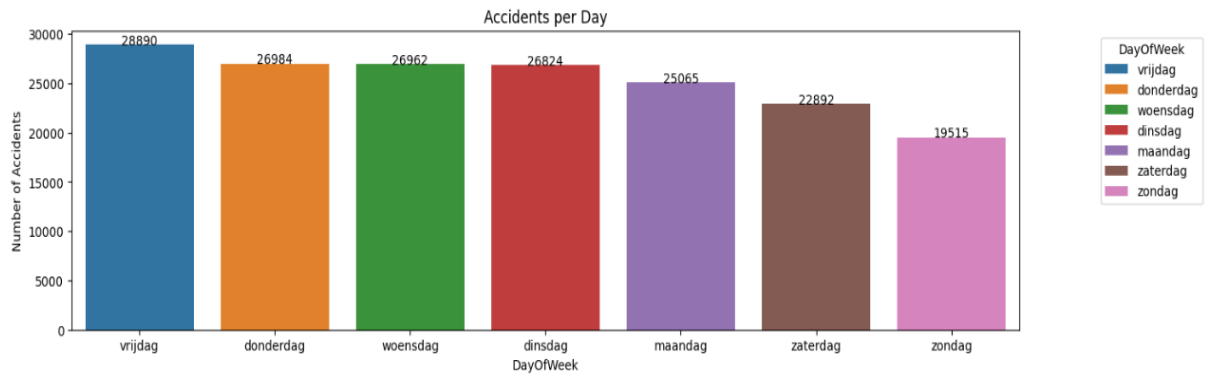
Bij een uurgranulaire analyse van de ongevallen blijkt het volgende:

- **Ochtendpieken:** Rond 8:00 en 9:00 uur is er een duidelijke piek te zien. *Interpretatie:* Ochtendspits veroorzaakt vaak extra verkeer omdat mensen haast hebben om op tijd op hun werk, school of andere verplichtingen te komen. Dit verhoogt de kans op ongevallen.
- **Avondpieken:** De meeste ongevallen vinden plaats tussen 16:00 en 19:00 uur. *Interpretatie:* In de avond is er weer een sterke verkeersstroom doordat mensen onderweg zijn van werk naar huis of kinderen ophalen. Deze drukte verhoogt eveneens de kans op incidenten.

De data toont dus aan dat de peaktijden samenhangen met de dagelijkse routines, zoals reizen tijdens de spitsuren. Dit leidt tot meer voertuigen op de weg en daardoor tot een hoger aantal ongevallen.



3.2. Dagelijke Ongevalanalyse:

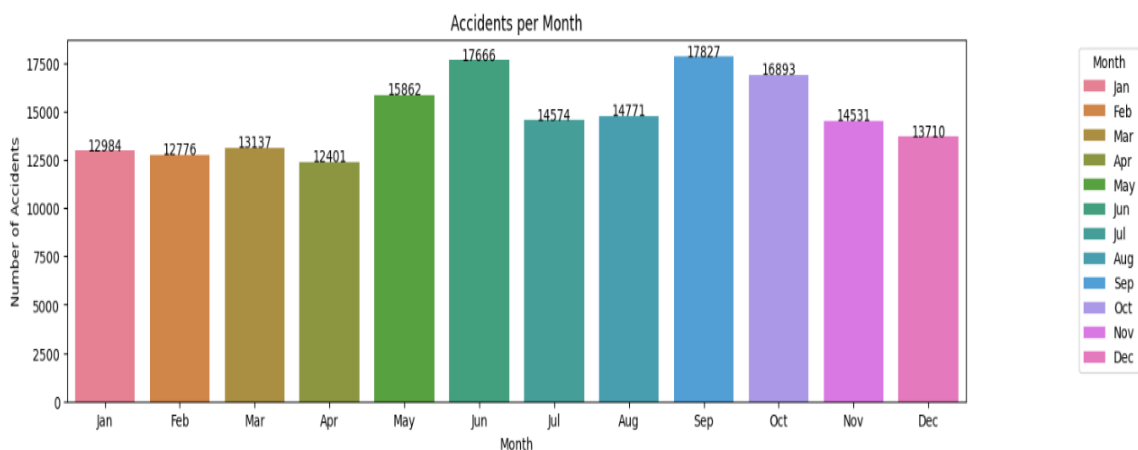


Hoogste ongevallen op vrijdag: De data toont aan dat de meeste ongevallen op vrijdag plaatsvinden, met een totaal van 28.890 ongevallen. Dit kan worden toegeschreven aan de einde-weekspits, wanneer mensen ernaar streven hun werkweek af te sluiten en met hun weekendactiviteiten te beginnen.

Laagste ongevallen op zondag: Zondagen kennen het laagste aantal ongevallen, namelijk 19.515 incidenten. Dit komt waarschijnlijk door het lichtere verkeer en het feit dat minder mensen naar werk of school gaan.

Weekdagen versus weekend: Over het algemeen liggen de ongevallen tijdens weekdagen hoger dan in het weekend. Dit sluit aan bij typische woon-werkverkeerspatronen, waarbij er tijdens weekdagen meer voertuigen op de weg zijn vanwege werk- en schoolverkeer.

3.3. Maandelijks Ongevallenanalyse

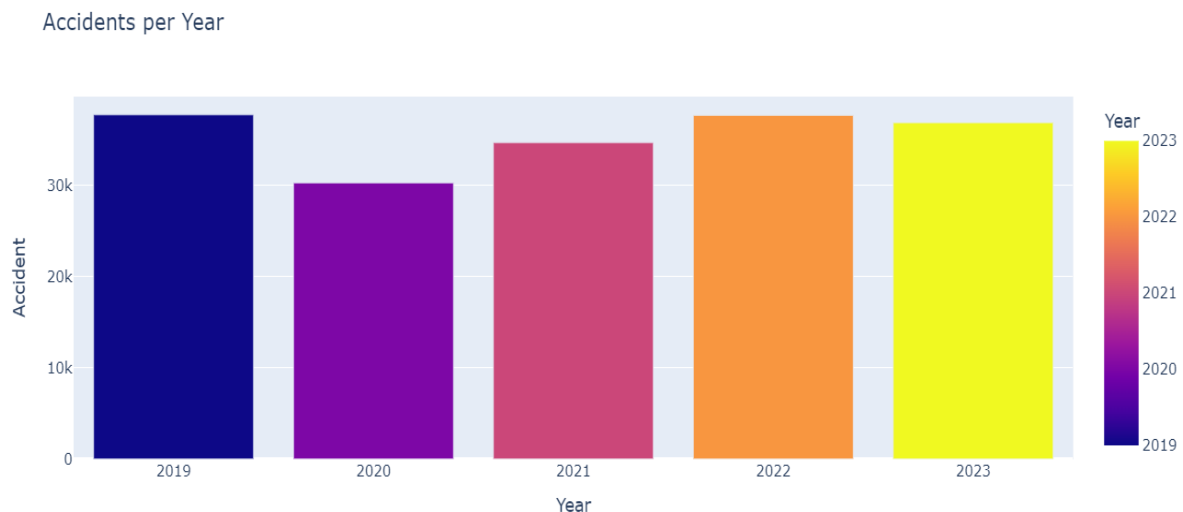


- **Hoogste ongevallen in september:** In september komen de hoogste aantallen ongevallen voor, namelijk 17.827 incidenten. Dit fenomeen kan worden verklaard door de terugkeer naar de normale dagelijkse routines na de zomervakanties. Mensen hervatten hun gebruikelijke activiteiten, waardoor het verkeer toeneemt en zo de kans op ongevallen stijgt.
- **Piekmomenten in juni en oktober:** Juni registreert 17.666 ongevallen en oktober 16.893 ongevallen. Deze relatief hoge cijfers kunnen samenhangen met specifieke

weersomstandigheden of een toename van reisactiviteiten in deze maanden. Mogelijk passen de lokale omstandigheden in deze periodes ongunstig in het rijgedrag of zorgt een verhoogde verkeersintensiteit voor meer incidenten.

- **Lente en zomer:** Er is een duidelijke stijging in het aantal ongevallen gedurende de late lente en zomermaanden (mei, juni, juli en augustus). Mogelijke verklaringen hiervoor zijn:
 - **Meer buitenshuis activiteiten en vakanties:** Mensen zijn actiever en maken vaker lange ritten tijdens de zomer, wat leidt tot een grotere blootstelling op de weg.
 - **Betere weersomstandigheden:** Hoewel het weer doorgaans gunstiger is, kan de combinatie van meer verkeer en hoge snelheden toch voor een toename van ongevallen zorgen.
- **Wintermaanden:** De cijfers in de winter (januari, februari en december) liggen lager in vergelijking met de pieken in de late lente en zomer. Desalniettemin blijven de ongevallen aantallen significant. Mogelijke verklaringen zijn:
 - **Voorzichtiger rijgedrag:** In de winter zijn bestuurders zich meer bewust van de gevaarlijke omstandigheden, zoals kou, sneeuw, ijs en slecht zicht, wat kan leiden tot een behoedzamer rijgedrag.
 - **Hinderlijke weersomstandigheden:** Hoewel men voorzichtiger rijdt, zorgen sneeuw en ijs er vaak voor dat de rijomstandigheden slechter zijn, wat alsnog tot incidentele ongevallen kan leiden.

3.4. Jaarlijkse trends in ongevallen



Trendanalyse:

2019: Het hoogste aantal ongelukken, wat aangeeft dat dit jaar mogelijk minder veiligheidsmaatregelen of hogere verkeersdrukke had.

2020: Een significante daling in het aantal ongelukken. Dit kan verband houden met beperkingen zoals lockdowns of verminderde mobiliteit als gevolg van de COVID-19-pandemie.

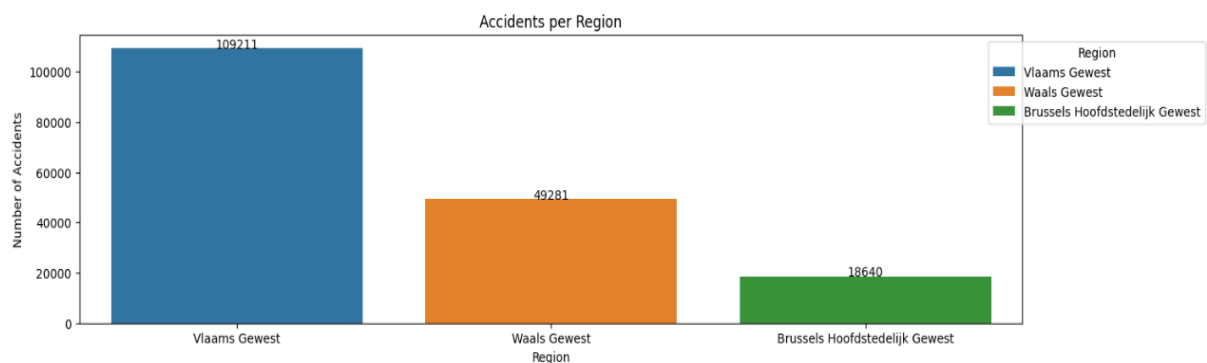
2021: Een lichte stijging, wat een gedeeltelijk herstel kan suggereren van het normale verkeer na de pandemie.

2022: Een duidelijke toename van ongevallen, wat wijst op mogelijk hogere verkeersdrukte en een mogelijke afname van waakzaamheid op de weg.

2023: Voortdurende toename, wat een aandachtspunt kan zijn voor beleidsmakers om verkeersveiligheid te verbeteren.

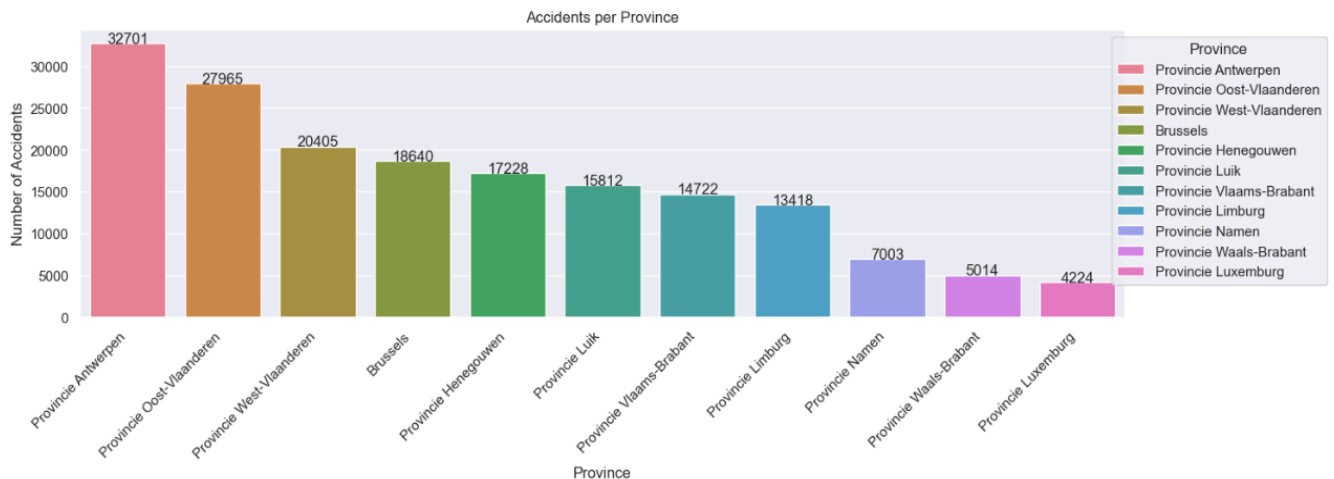
4. Regional Ongeval analyse:

4.1. Gewestelijk analyse:



Het Vlaams Gewest registreert het hoogste aantal ongevallen en verwondingen, wat de uitgebreide wegeninfrastructuur en de hoge bevolkingsdichtheid weerspiegelt. Ter vergelijking heeft het Brussels Hoofdstedelijk Gewest de laagste cijfers, waarschijnlijk dankzij effectief stedelijk verkeersmanagement.

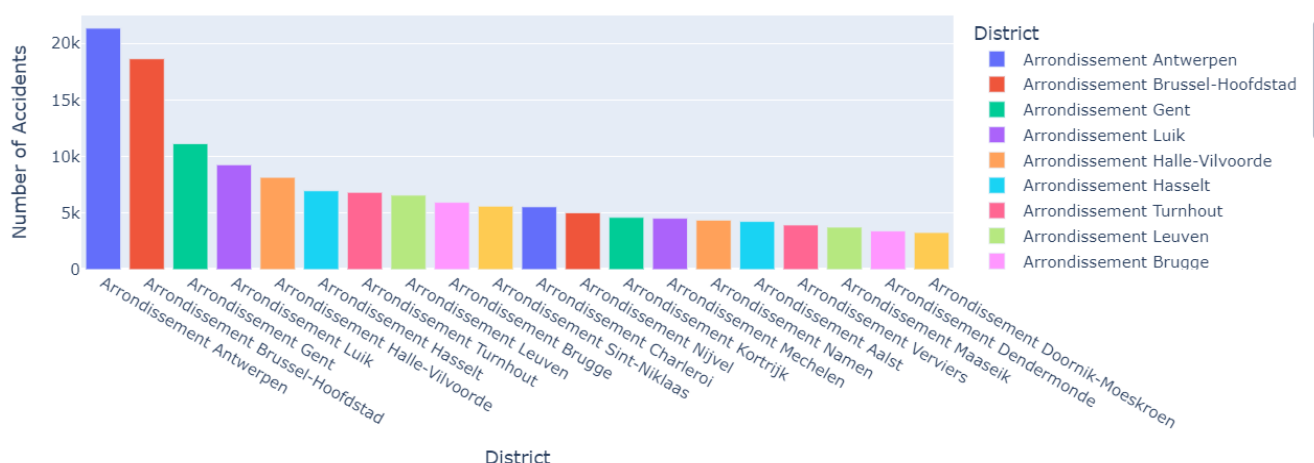
4.2. Provinciaal analyse:



- **Antwerpen:** Hoogste totaal aantal ongevallen (32.701) en lichte verwondingen (29.386), wat wijst op een zeer intensieve verkeersactiviteit.
- **Brussel:** Een aanzienlijk aantal ongevallen (18.640), maar relatief weinig dodelijke ongevallen (45) en ernstige verwondingen (795), waarschijnlijk door het effectieve stedelijke verkeersmanagement.
- **Henegouwen:** Een hoog sterftecijfer (307) en een groot aantal ernstige verwondingen (1.367), wat de noodzaak voor veiligheidsverbeteringen benadrukt.
- **Andere provincies:** Matige tot hoge ongevallen- en verwondingscijfers, waarbij bepaalde provincies hogere percentages van dodelijke ongevallen en ernstige verwondingen vertonen, wat mogelijk gerichte veiligheidsmaatregelen vereist.

4.3. Gemeentelijke analyses

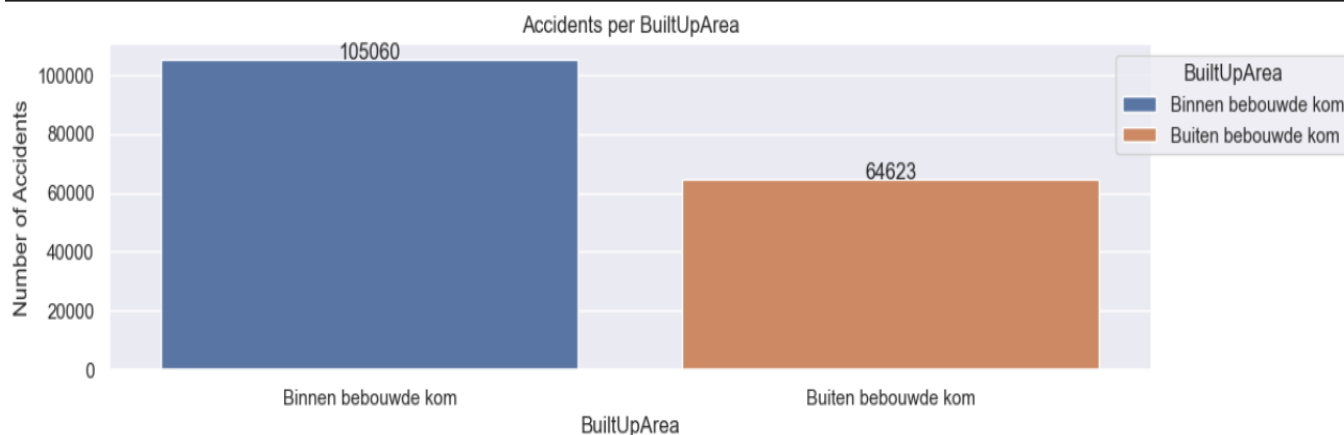
Accidents per District



De hoge aantallen in Antwerpen en Brussel kunnen te wijten zijn aan hogere verkeersdruk, stedelijke gebieden met meer voertuigen, of complexe verkeerssituaties.

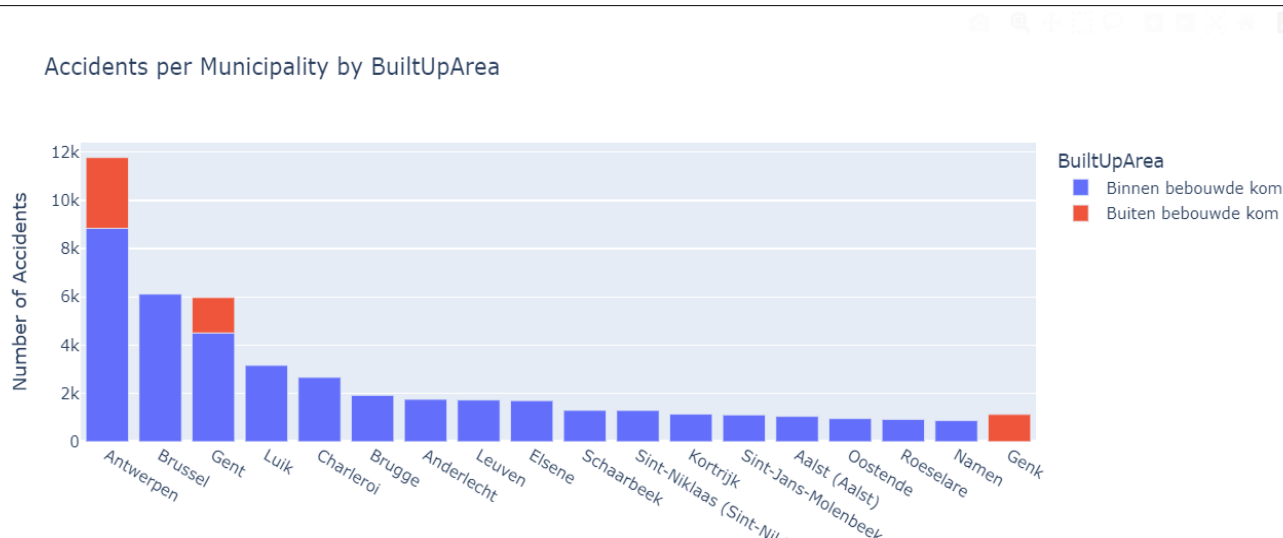
De lagere aantallen in kleinere arrondissementen zoals Veurne of Neufchâteau kunnen samenhangen met minder verkeer en minder bevolkingsdichtheid.

4.4. Analyse van ongevallen in bebouwde versus niet-bebouwde gebieden



Meer ongevallen in bebouwde gebieden: In de meeste provincies is het aantal ongevallen hoger in bebouwde gebieden vergeleken met de gebieden buiten de bebouwde omgeving. Deze trend is vooral duidelijk in stedelijke regio's zoals Brussel en de Provincie Antwerpen.

Significant verschil in Provincie Limburg: Provincie Limburg vormt een uitzondering: daar vinden meer ongevallen plaats buiten bebouwde gebieden (8.336) dan binnen bebouwde gebieden (4.582). Dit duidt erop dat verkeersveiligheidskwesties in landelijke of minder dichtbevolkte gebieden een belangrijke zorg zijn in Limburg.

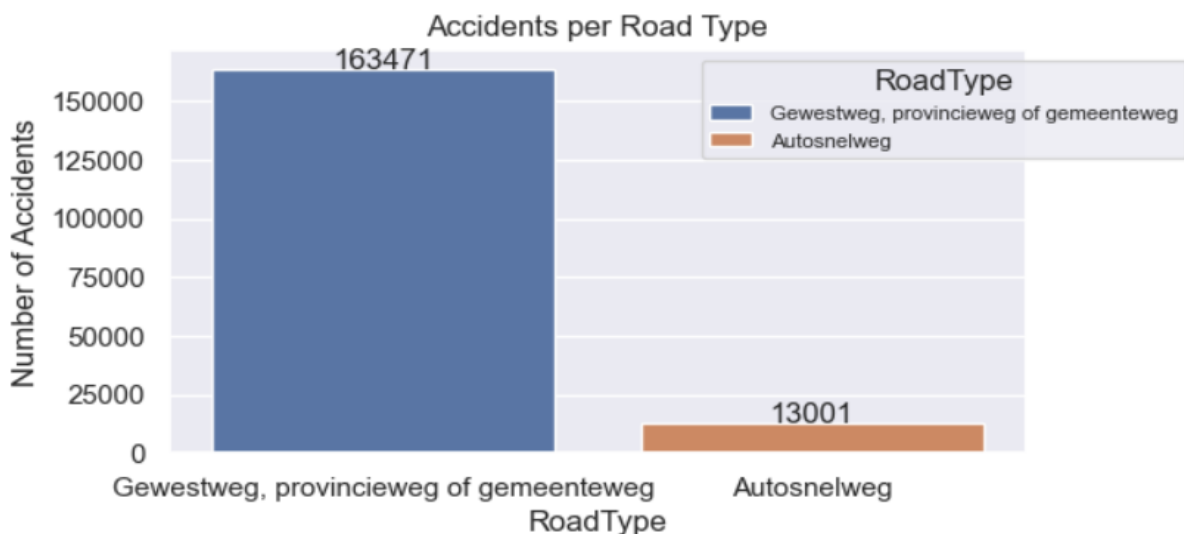


Stedelijk versus landelijk contrast: In dichtbevolkte provincies zoals Brussel en Antwerpen vindt de meerderheid van de ongevallen binnen bebouwde gebieden plaats, wat te wijten is aan de hogere verkeersdichtheid, meer kruispunten en grotere voetgangersactiviteit. In landelijke provincies zoals Luxemburg en Waals-Brabant is het totale aantal ongevallen lager, maar de verdeling tussen bebouwde en niet-bebouwde gebieden is relatief in balans.

Provincies met veel ongevallen in bebouwde gebieden: De Provincie Antwerpen (18.634 ongevallen) en de Provincie Oost-Vlaanderen (16.246 ongevallen) noteren hoge aantallen ongevallen in bebouwde gebieden, wat het belang onderstreept van gerichte maatregelen voor verkeersveiligheid in stedelijke omgevingen.

5. Wegtype Analyse

5.1. Snelwegen versus hoofdwegen



Autosnelweg (Highway)

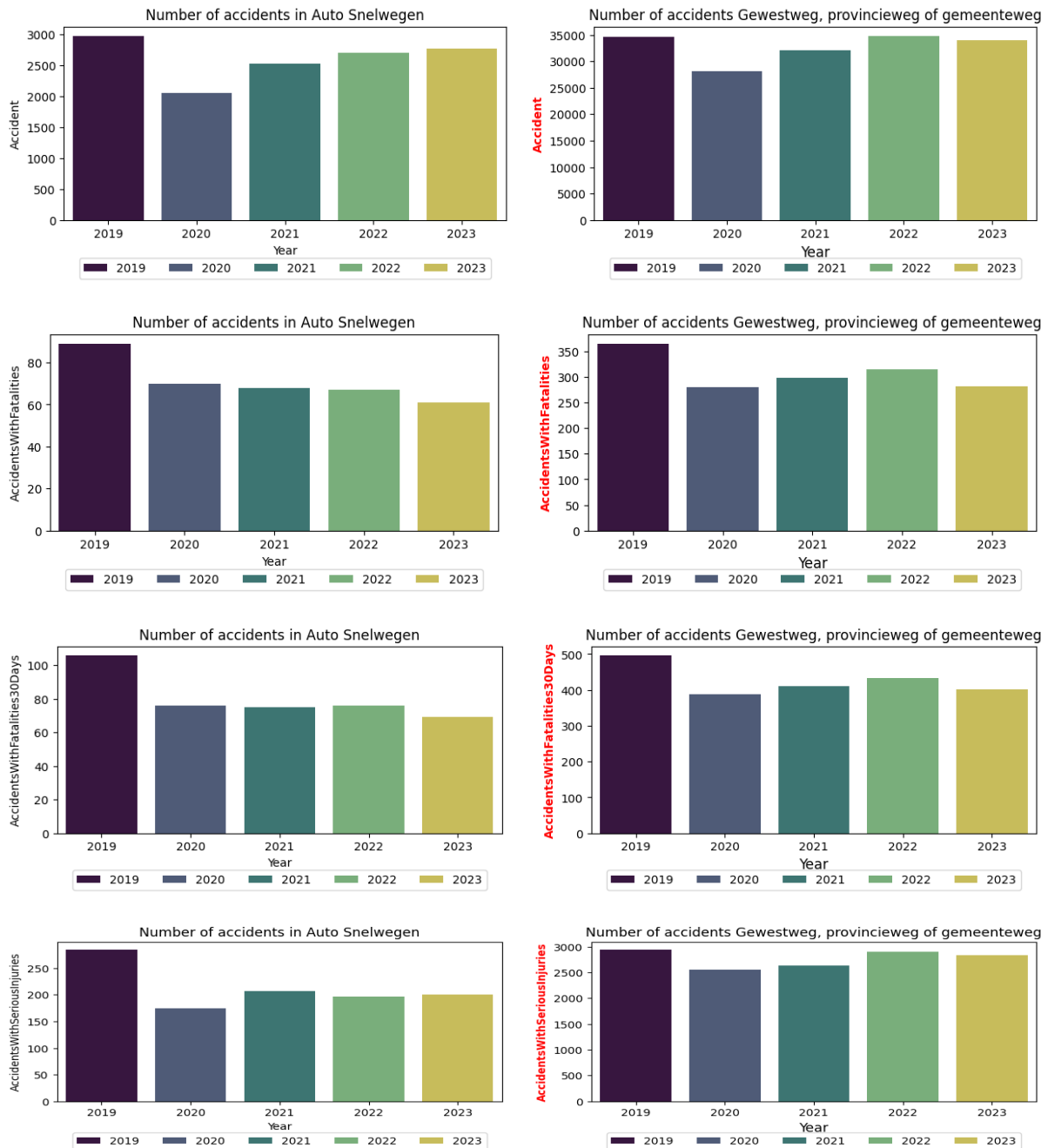
- **Ongevallen en letselcijfers:** Op autosnelwegen vindt een aanzienlijk aantal ongevallen plaats (13.001 ongevallen).
- **Dodelijke ongevallen:** Ondanks het hoge totaal aantal ongevallen is het aantal dodelijke ongevallen relatief laag (355).
- **Ernstige en lichte verwondingen:** Ook het aantal ernstige en lichte verwondingen op snelwegen is aanzienlijk. Dit toont aan dat hoge snelheden – ondanks een relatief lager sterftecijfer – leiden tot impacten met ernstige gevolgen.

Gewestweg, Provincieweg of Gemeenteweg (Regional, Provincial, or Municipal Roads)

- **Ongevallen en gebruik:** Deze wegen vertegenwoordigen het hoogste totaal aantal ongevallen (163.471). Dit is waarschijnlijk terug te voeren op het uitgebreide netwerk en het hoge gebruik door zowel lokaal als doorrijdend verkeer.
- **Letselcijfers:** Ze laten ook de hoogste cijfers zien in andere categorieën, zoals:
 - Dodelijke ongevallen: 1.539
 - Dodelijke ongevallen binnen 30 dagen: 2.130
 - Ernstige verwondingen: 13.864

- Lichte verwondingen: 147.477
- Kleinere verwondingen: 591

5.2. Overzicht van ongevallen, ernstige verwondingen en fataliteiten



5.3. Risico's en verkeersintensiteit op verschillende wegtypen

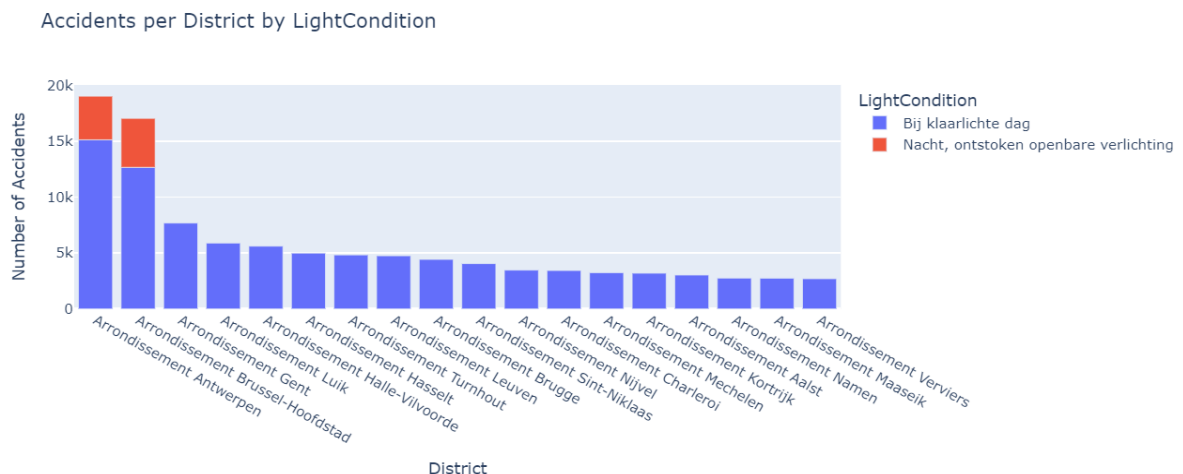
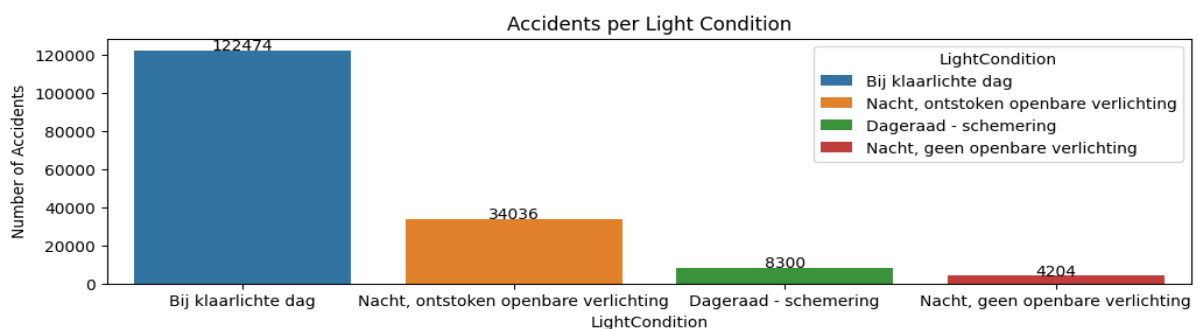
Autosnelwegen (Highways) versus Hoofdwegen (Main Roads): Autosnelwegen laten een aanzienlijk lager totaal aantal ongevallen zien vergeleken met hoofdwegen. Toch benadrukt de relatief consistente hoeveelheid ernstige verwondingen en dodelijke ongevallen dat er op

autosnelwegen geen vertrouwen is op een lagere risico-indicatie. Dit onderstreept de noodzaak van voortdurende veiligheidsmaatregelen, ook op autosnelwegen.

Ernstige Verwondingen: Zowel autosnelwegen als hoofdwegen registreren een aanzienlijk aantal ernstige verwondingen. Dit benadrukt de potentieel ernstige gevolgen van verkeersongevallen op beide wegtypen. Vooral op autostrades, waar het aantal ongevallen hoger ligt, komen van nature ook meer ernstige verwondingen voor.

Trends in Dodelijkheid: Het aantal dodelijke ongevallen binnen 30 dagen na een ongeval ligt iets hoger dan het aantal onmiddellijke dodelijke ongevallen. Dit wijst op het cruciale belang van tijdige medische interventies en effectieve veiligheidsmaatregelen om de gevolgen van ongevallen te beperken.

6. Analyse van Lichtcondities

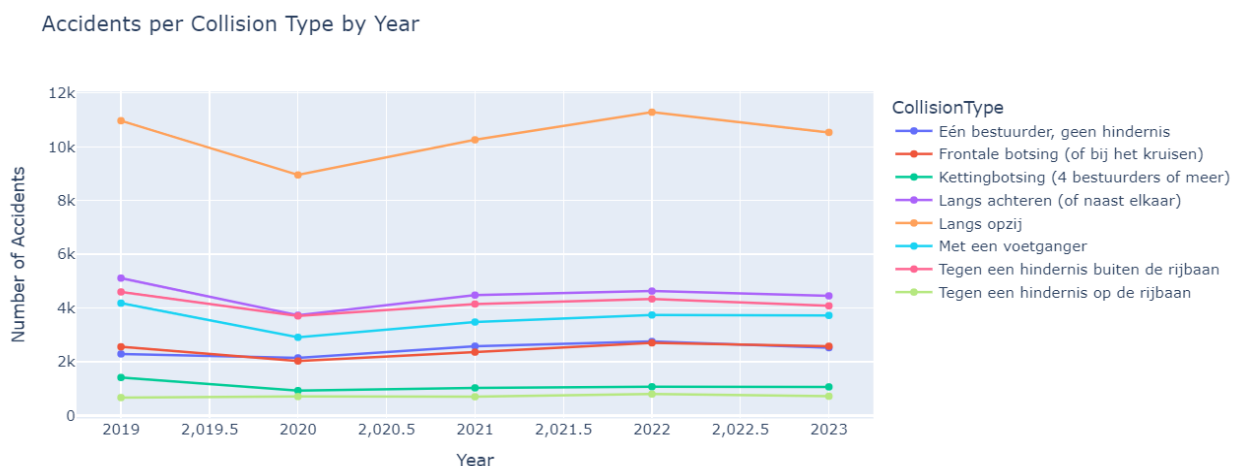
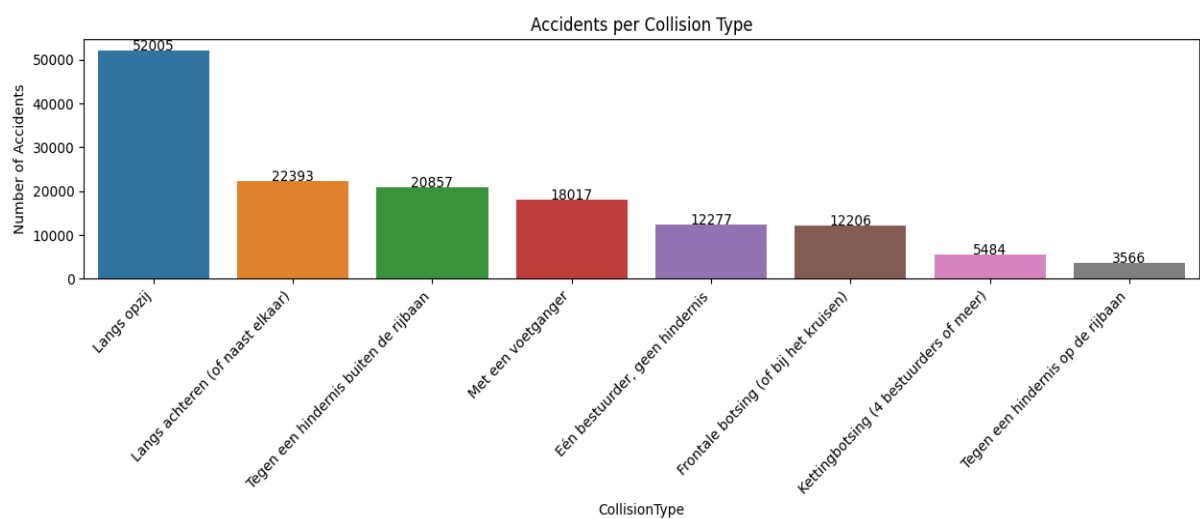


- **Daglichtcondities:** De meeste ongevallen vinden plaats overdag. De hoogste aantallen zijn te zien in Provincie Antwerpen, Provincie Oost-Vlaanderen en Provincie West-Vlaanderen.
- **Schemeringscondities (dageraad en schemering):** Hoewel het aantal ongevallen tijdens de overgang tussen dag en nacht (dageraad en schemering) lager ligt dan overdag, is het aantal toch significant. Provincie Antwerpen en Provincie Oost-Vlaanderen lopen in deze categorie voorop.

- **Nacht zonder openbare verlichting:** Tijdens de nacht zonder openbare verlichting treden over het algemeen minder ongevallen op. Wel rapporteren Provincie Oost-Vlaanderen en Provincie West-Vlaanderen de hoogste aantallen in deze categorie.
- **Nacht met openbare verlichting:** Een aanzienlijk aantal ongevallen gebeurt 's nachts wanneer er wel openbare verlichting aanwezig is. Hierbij treden de hoogste aantallen op in Provincie Antwerpen, Provincie Oost-Vlaanderen en in het Brussels Hoofdstedelijk Gewest.

7. Analyse van Botsingstypen

7.1. Overzicht en trends per botsingstype:



Het jaar 2020 kende een duidelijke vermindering van het aantal ongevallen voor de meeste typen, waarschijnlijk als gevolg van minder verkeer tijdens de COVID-19-pandemie. Sindsdien zijn de cijfers echter weer gestegen, wat de aanhoudende aard van deze problemen onderstreept.

- **Side Impact (Langs opzij):** Dit is het meest voorkomende type botsing, dat vaak voorkomt op kruispunten.

- **Rear-end (Langs achteren):** Veelvoorkomend bij druk verkeer of bij stoplichten.
- **With Pedestrian (Met een voetganger):** Een significant aantal botsingen, met name in stedelijke gebieden.
- **Against an Obstacle (Tegen een hindernis):** Zowel op als naast de weg wijzen deze botsingen op problemen met het wegontwerp of op afleiding van de bestuurder.

8. Zakelijke vragen en mogelijke antwoorden:

Op basis van de bovenstaande analyses zijn we nu in staat om onze Businessvragen diepgaand te beantwoorden:

8.1. Welke factoren hebben de grootste invloed op het aantal en de ernst van verkeersongevallen in België?

Er zijn verschillende factoren geïdentificeerd die zowel het aantal als de ernst van verkeersongevallen sterk beïnvloeden:

- **Tijdstip en Dagelijkse Ritmes:**
 - **Uurlijkse Verdeling:** De data toont duidelijke pieken in zieken tussen 08:00–09:00 en 16:00–19:00 uur. Deze piekmomenten vallen samen met de ochtend- en avondspits, waarin het verkeer heviger is en bestuurders sneller moeten reageren in drukke situaties.
 - **Dagelijkse Verdeling:** Vrijdagen hebben het hoogste aantal ongevallen, wat mogelijk wordt veroorzaakt door de einde-weekspits wanneer mensen snel hun werkweek willen afronden en hun weekend beginnen. Daarentegen zijn zondagen het rustigst, met het laagste aantal ongevallen.
- **Maandelijks en Jaarrapportages:**
 - **Seizoensinvloeden:** Er is een sterke seizoensgebonden trend zichtbaar. September, juni en oktober registreren hoge ongevallen aantallen, wat samenhangt met het toegenomen weggebruik tijdens de overgang van zomer naar herfst. Ook de late lente en de zomermaanden (mei t/m augustus) vertonen hogere cijfers, waarschijnlijk door meer buitenactiviteiten, vakanties en dus meer verkeer.
 - **COVID-19 Effecten:** Het jaar 2020 kende een opvallende daling in het aantal ongevallen door de beperkende maatregelen, waarna de cijfers in de daaropvolgende jaren weer omhoog gingen. Dit benadrukt dat verkeersintensiteit en -gedrag cruciale variabelen zijn.
- **Wegtype en Omgevingsfactoren:**
 - **Type Weeg:** Ongevallen op regionale, provinciale en gemeentelijke wegen zijn talrijker dan op autosnelwegen. Dit is deels te wijten aan de complexiteit van deze wegen – een uitgebreid netwerk, gemengd verkeer, wisselende wegkwaliteit en minder strikte veiligheidsmaatregelen vergeleken met hoogveiligheidsautosnelwegen.

- Bebouwde versus Landelijke Gebieden: In de meeste provincies komen ongevallen vaker voor in bebouwde gebieden. Een uitzondering is Provincie Limburg, waar juist meer ongevallen buiten de bebouwde omgeving plaatsvinden. Dit wijst op specifieke risico's in landelijke of minder bevolkte gebieden.
- Lichtcondities:
 - Ongevallen variëren sterk afhankelijk van de lichtomstandigheden. De meeste incidenten vinden overdag plaats, maar ook bij schemering (dageraad en schemering) zien we aanzienlijke aantallen, wat suggereert dat overgangen in lichtcondities bestuurders voor uitdagingen stellen.
 - Nachtelijke ongevallen verschillen tussen situaties met en zonder openbare verlichting, wat wijst op het belang van infrastructuur en zichtbaarheid.
- Botsingstypen en Impact:
 - Verschillende botsingstypen spelen een rol. Side impact (langs opzij) is het meest frequent, vooral bij kruispunten, terwijl achteraanrijdingen veel voorkomen bij druk verkeer of bij stoplichten.
 - Ook botsingen met voetgangers en tegen obstakels zijn belangrijk, omdat deze type incidenten vaak leiden tot ernstige verwondingen en duiden op aandachtspunten zoals wegontwerp en afleiding van bestuurders.
- Regionale Demografie en Infrastructuur:
 - Een hogere dichtheid van inwoners en een uitgebreid wegennet (zoals in het Vlaams Gewest) leiden tot een hoger aantal incidenten. Tegelijkertijd laat de analyse zien dat effectief verkeersmanagement (zoals in het Brussels Hoofdstedelijk Gewest) kan bijdragen aan lagere cijfers van ernstige ongevallen en fataliteiten.

Kortom, een combinatie van verkeersintensiteit, tijdsgebonden patronen, wegtype, licht- en weersomstandigheden, en regionale factoren (infrastructuur en demografie) bepalen samen het aantal en de ernst van verkeersongevallen in België.

8.2. Hoe verschillen verkeersongevallen per regio en tijdstip, en zijn er bepaalde trends te ontdekken?

De data-analyse wijst op duidelijke verschillen en trends in verkeersongevallen, afhankelijk van regio en tijdstip:

- Tijdstipvariaties:
 - Uurlijkse Verdeling: Ongevallen pieken 's ochtends (rond 08:00–09:00) en in de namiddag/avond (16:00–19:00), wat voornamelijk samenvalt met de spitsuren. Dit benadrukt het effect van verhoogde verkeersdichtheid tijdens de woon-werkspits.
 - Dagelijkse Verdeling: Vrijdagen hebben de hoogste aantallen ongevallen, wat waarschijnlijk bijdraagt aan de “einde-weekspits” en het veranderende rijgedrag

tegen het einde van de werkweek. Arteriële afnames worden waargenomen op zondagen, wanneer het verkeer rustig is.

- Maandelijks en Jaarlijkse Trends: September registreert de hoogste ongevallen, terwijl april de laagste cijfers hebben. Daarnaast zorgen seizoensveranderingen (meer ongevallen in de late lente en zomer) en de invloed van bijzondere omstandigheden (zoals de COVID-19 gerelateerde daling in 2020) voor herkenbare trends.
-
- Regionale Verschillen:
 - Gewest en Provincieverschillen: Het Vlaams Gewest registreert het hoogste aantal ongevallen en letsels, hetgeen verklaard kan worden door een grotere populatie en een uitgebreider wegennet. Het Brussels Hoofdstedelijk Gewest toont lagere cijfers, wat wijst op de effectiviteit van stedelijk verkeersmanagement. Ook op provinciaal niveau zien we verschillen: Antwerpen en Oost-Vlaanderen scoren erg hoog, vooral in bebouwde gebieden, terwijl provincies zoals Limburg juist meer ongevallen buiten de bebouwde omgeving ervaren.
 - Stedelijke versus Landelijke Omgeving: In dichtbevolkte stedelijke gebieden komen meer ongevallen voor, wat verband houdt met hogere verkeersdichtheden, meer kruispunten en intensere voetgangersactiviteit. In landelijke gebieden, hoewel het totale ongevallen aantal vaak lager is, kunnen de ongevallenstemperaturen verschillend zijn door andere factoren als snelheid en wegontwerp.
- Trend Observaties:
 - Er is een duidelijke trend zichtbaar in verband met de dagelijkse, wekelijkse en maandelijkse distributie van ongevallen, wat de invloed van veranderende verkeerspatronen en gebruikspatronen benadrukt.
 - Regionale analyses tonen dat gebieden met een hoge bevolkingsdichtheid en veel verkeersintensiteit (zoals Antwerpen) speciale aandacht vereisen, terwijl ook de verschillen tussen stedelijke en landelijke gebieden om gerichte maatregelen vragen.

9. Samenvatting en Conclusies

Samenvattend laten de analyses zien dat zowel het tijdstip van de dag als de regionale context een grote rol spelen bij het aantal en de ernst van verkeersongevallen. Deze inzichten bieden handvatten om gerichte verkeersveiligheidsmaatregelen te ontwikkelen, afgestemd op zowel de specifieke risico's tijdens piekmomenten als op de regionale verschillen in infrastructuur en verkeersbeheer.

Dit project biedt waardevolle inzichten in de factoren die verkeersongevallen in België beïnvloeden. Door te focussen op risicovolle gebieden en gerichte veiligheidsmaatregelen te implementeren, kunnen we werken aan een verlaging van zowel het aantal als de ernst van ongevallen. Uiteindelijk leidt dit tot veiligere wegen voor iedereen.

