Pedal Cycle Traffic Report

In this project, we will analyze the pedal cycle traffic data, provided by the Department for Transport, which is segmented by road classes, regions, and vehicle types. However, our specific focus in this report will be on pedal traffic during the period from 1993 to 2022. This analysis aims to shed light on locations and years with the greatest and least impact, examining how events such as COVID-19 have brought about changes in pedal traffic patterns.

Analyzing pedal traffic data is highly beneficial, offering insights into urban sustainability, public health, and transportation planning. The data provides valuable information for urban authorities to enhance cycling infrastructure, improve road safety, and adapt policies during events like the COVID-19 pandemic. Beyond its positive impact on environmental sustainability, increased cycling contributes to economic growth, promotes healthier lifestyles, and fosters more accessible and efficient cities. The analysis of pedal traffic patterns emerges as a crucial tool for shaping policies that advance urban development, prioritize public health, and create thriving, sustainable communities.

Dataset

The chosen dataset is the workbook TRA0413 collected from official UK government website, produced by Department of Transport in the collection Road traffic statistics (TRA). It describe the pedal cycle traffic (vehicle kilometers) by region and county in Great Britain, annual from 1993. Being updated annually, at this time the most recent data refers to 2022.

- **Quality**: The data is constantly updated, with a wide variety of datasets focusing on different issues. Being used by the department for real life applications, giving us confidence in use it. In addition to taking care of attach notes to the data to provide more information. It shows the good quality of data.

- **Detail**: The data consists on pedal traffic on Great Britain by regions, counties over the years, and some notes is available in dataset to notify about impacts on traffic volume in some years. The traffic volume is measured by kilometers traveled by vehicle.

- **Documentation**: The Road traffic statistics (TRA) collection offers a good documentation with notes, definitions, reviews, reports and methodologies used. Available on page of Road traffic statistics information in official United Kingdom government website.

- **Interrelation**: We could use dataset with another datasets with cars or general vehicles information or with traffic locations and road classes for a better understanding of the situation and traffic and its changes. however, we would not have enough information to link the data, for example in the case of road classes and their locations.

- **Use**: The purpose of using this dataset is to better understand pedal traffic for shaping policies that advance urban development, prioritize public health, and create thriving, sustainable communities. In addition to analyzing the impacts of events on pedal traffic.

- **Discoverability**: It may not be so easy to find the specific dataset due to the number of datasets and collections available on website, but with the code tra0413 it is not too difficult to find it.

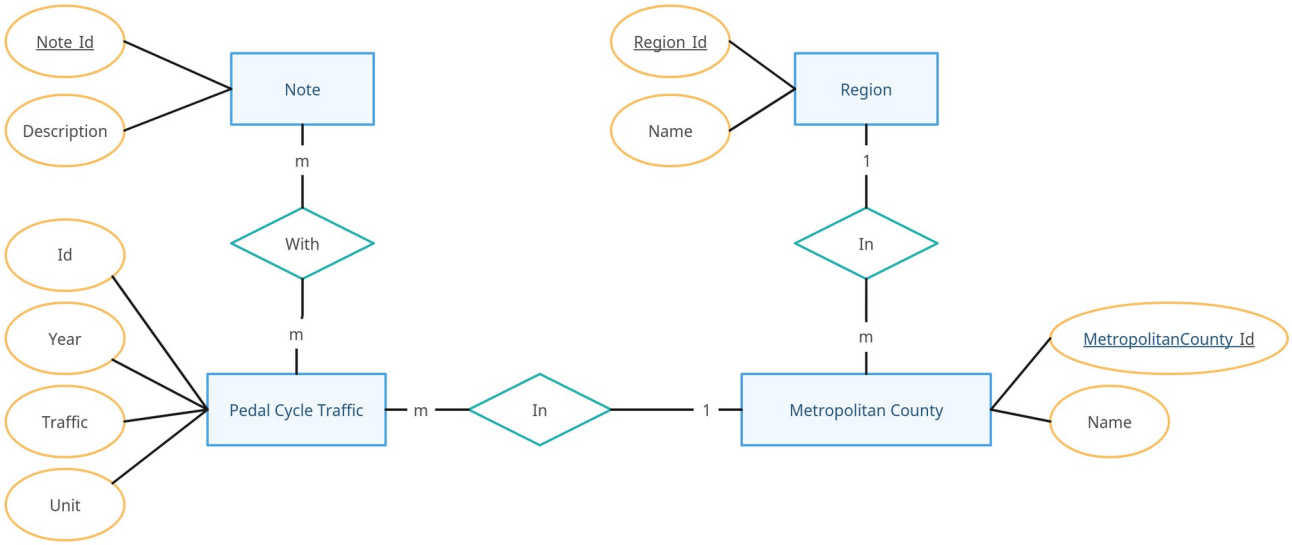
Useful links

Below are some useful links to find the dataset used and its documentation:

- Dataset: https://assets.publishing.service.gov.uk/media/64af0c898bc29f00132ccc8a/tra0413.ods
- Collection: https://www.gov.uk/government/collections/road-traffic-statistics#latest-road-traffic-statistics
- Documentation: https://www.gov.uk/guidance/road-traffic-statistics-information#general-road-traffic-statistics-information

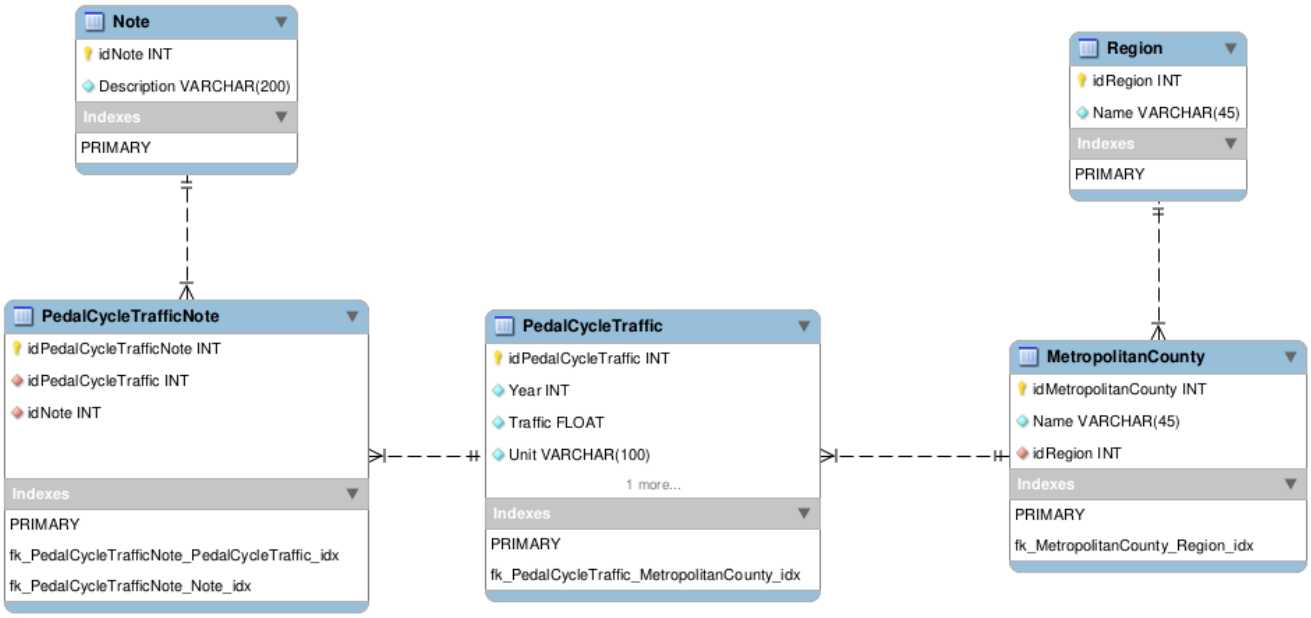
ER Model

Entity–relationship model created from dataset:



Diagram

Entity–relationship Diagram created from ER model in previous step:



Building database

Below is the SQL code for creating the schema and tables discussed previously in the modeling stage:

```
CREATE SCHEMA IF NOT EXISTS `tra0413` ;
USE `tra0413`;
```

```
+-----------------------+
| Tables_in_tra0413     |
+-----------------------+
| MetropolitanCounty    |
| Note                  |
| PedalCycleTraffic     |
| PedalCycleTrafficNote |
| Region                |
+-----------------------+
```

This is the result of how it should look after all tables are created

Regions

```
CREATE TABLE IF NOT EXISTS `tra0413`.`Region` (
  `idRegion` INT NOT NULL,
  `Name` VARCHAR(45) NOT NULL,
  PRIMARY KEY (`idRegion`))
ENGINE = InnoDB;
```

```
+----------+-------------+------+-----+---------+-------+
| Field    | Type        | Null | Key | Default | Extra |
+----------+-------------+------+-----+---------+-------+
| idRegion | int         | NO   | PRI | NULL    |       |
| Name     | varchar(45) | NO   |     | NULL    |       |
+----------+-------------+------+-----+---------+-------+
```

Notes

```
CREATE TABLE IF NOT EXISTS `tra0413`.`Note` (
  `idNote` INT NOT NULL,
  `Description` VARCHAR(200) NOT NULL,
  PRIMARY KEY (`idNote`))
ENGINE = InnoDB;
```

```
+-------------+--------------+------+-----+---------+-------+
| Field       | Type         | Null | Key | Default | Extra |
+-------------+--------------+------+-----+---------+-------+
| idNote      | int          | NO   | PRI | NULL    |       |
| Description | varchar(200) | NO   |     | NULL    |       |
+-------------+--------------+------+-----+---------+-------+
```

Counties

```
CREATE TABLE IF NOT EXISTS `tra0413`.`MetropolitanCounty` (
  `idMetropolitanCounty` INT NOT NULL,
  `Name` VARCHAR(45) NOT NULL,
  `idRegion` INT NOT NULL,
  PRIMARY KEY (`idMetropolitanCounty`),
  INDEX `fk_MetropolitanCounty_Region_idx` (`idRegion` ASC) VISIBLE,
  CONSTRAINT `fk_Metropolitan County_Region`
    FOREIGN KEY (`idRegion`)
    REFERENCES `tra0413`.`Region` (`idRegion`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION)
ENGINE = InnoDB;
```

```
+---------------------+-------------+------+-----+---------+-------+
| Field               | Type        | Null | Key | Default | Extra |
+---------------------+-------------+------+-----+---------+-------+
| idMetropolitanCounty | int         | NO   | PRI | NULL    |       |
| Name                | varchar(45) | NO   |     | NULL    |       |
| idRegion            | int         | NO   | MUL | NULL    |       |
+---------------------+-------------+------+-----+---------+-------+
```

Pedal Traffic

```
CREATE TABLE IF NOT EXISTS `tra0413`.`PedalCycleTraffic` (
  `idPedalCycleTraffic` INT NOT NULL,
  `Year` INT NOT NULL,
  `Traffic` FLOAT NOT NULL,
  `Unit` VARCHAR(100) NOT NULL,
  `idMetropolitanCounty` INT NOT NULL,
  PRIMARY KEY (`idPedalCycleTraffic`),
  INDEX `fk_PedalCycleTraffic_MetropolitanCounty_idx` (`idMetropolitanCounty` ASC)
VISIBLE,
  CONSTRAINT `fk_Pedal Cycle Traffic_Metropolitan County`
    FOREIGN KEY (`idMetropolitanCounty`)
    REFERENCES `tra0413`.`MetropolitanCounty` (`idMetropolitanCounty`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION)
ENGINE = InnoDB;
```

Field	Type	Null	Key	Default	Extra
idPedalCycleTraffic	int	NO	PRI	NULL	
Year	int	NO		NULL	
Traffic	float	NO		NULL	
Unit	varchar(100)	NO		NULL	
idMetropolitanCounty	int	NO	MUL	NULL	

Pedal Traffic - Notes

```
CREATE TABLE IF NOT EXISTS `tra0413`.`PedalCycleTrafficNote` (
  `idPedalCycleTrafficNote` INT NOT NULL,
  `idPedalCycleTraffic` INT NOT NULL,
  `idNote` INT NOT NULL,
  PRIMARY KEY (`idPedalCycleTrafficNote`),
  INDEX `fk_PedalCycleTrafficNote_PedalCycleTraffic_idx` (`idPedalCycleTraffic`
ASC) VISIBLE,
  INDEX `fk_PedalCycleTrafficNote_Note_idx` (`idNote` ASC) VISIBLE,
  CONSTRAINT `fk_Pedal Cycle Traffic`
    FOREIGN KEY (`idPedalCycleTraffic`)
    REFERENCES `tra0413`.`PedalCycleTraffic` (`idPedalCycleTraffic`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION,
  CONSTRAINT `fk_Note`
    FOREIGN KEY (`idNote`)
    REFERENCES `tra0413`.`Note` (`idNote`)
    ON DELETE NO ACTION
    ON UPDATE NO ACTION)
ENGINE = InnoDB;
```

```
+--------------------------+------+------+-----+---------+-------+
| Field                    | Type | Null | Key | Default | Extra |
+--------------------------+------+------+-----+---------+-------+
| idPedalCycleTrafficNote | int  | NO   | PRI | NULL    |       |
| idPedalCycleTraffic     | int  | NO   | MUL | NULL    |       |
| idNote                   | int  | NO   | MUL | NULL    |       |
+--------------------------+------+------+-----+---------+-------+
```

Dataset reflection

The built database reflects the dataset well, but perhaps the application of the notes available in the dataset are not very useful for the questions we asked, just for additional information. However, it was kept in the database to keep the information linked.

Furthermore, the way the dataset was assembled, with each year representing a column of the .ods file, resulted in greater difficulty in processing the data and inserting it into the normalized database.

Queries

Below are the questions created based on the dataset and the queries to answer them

Which years had highest pedal traffic?

```sql
SELECT pct.`Year`, ROUND(SUM(pct.Traffic), 2) AS year_total
FROM tra0413.PedalCycleTraffic pct
GROUP BY pct.`Year`
ORDER BY year_total DESC;
```

Which regions had the highest pedal traffic in last year?

```sql
SELECT region_traffic.Name AS Region, ROUND(SUM(region_traffic.Traffic), 2) AS
total_traffic
FROM (
      SELECT r.Name, pct.Traffic
      FROM tra0413.PedalCycleTraffic pct
      LEFT JOIN tra0413.MetropolitanCounty mc
      ON pct.idMetropolitanCounty = mc.idMetropolitanCounty
      LEFT JOIN tra0413.Region r
      ON mc.idRegion = r.idRegion
      WHERE pct.`Year` = (
            SELECT MAX(pct.`Year`)
            FROM tra0413.PedalCycleTraffic pct
      )
) AS region_traffic
GROUP BY region_traffic.Name
ORDER BY total_traffic DESC;
```

Which county had the highest pedal traffics in last year?

```
SELECT
      county_traffic.Name AS County,
      ROUND(SUM(county_traffic.Traffic), 2) AS total_traffic
FROM(
      SELECT mc.Name, pct.Traffic  FROM tra0413.PedalCycleTraffic pct
      LEFT JOIN tra0413.MetropolitanCounty mc
      ON pct.idMetropolitanCounty = mc.idMetropolitanCounty
      WHERE pct.`Year` = (
            SELECT MAX(pct.`Year`)
            FROM tra0413.PedalCycleTraffic pct
      )
) AS county_traffic
GROUP BY county_traffic.Name
ORDER BY total_traffic DESC;
```

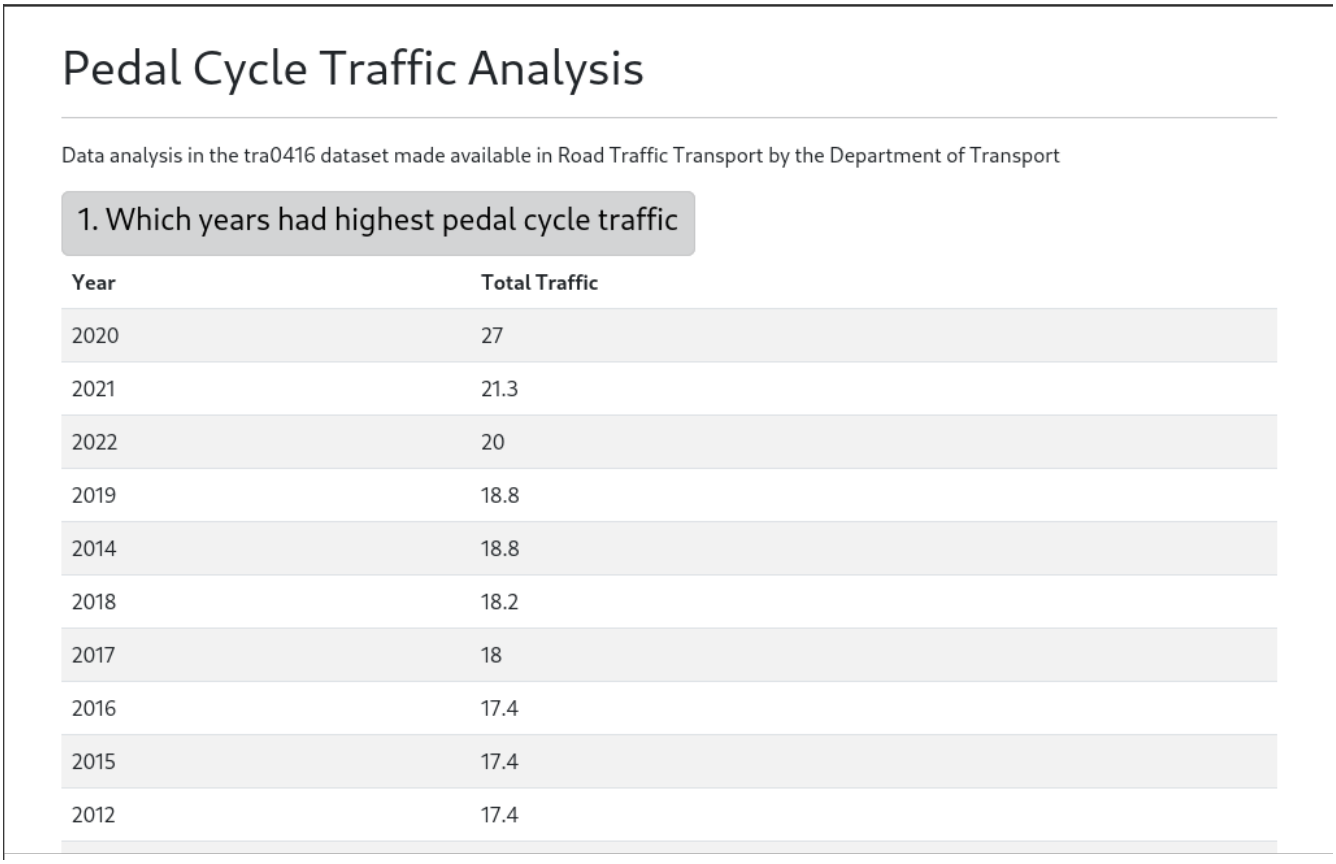
How much did COVID-19 affected pedal traffics?

```
SELECT
      traffic2019.Name AS County,
      ROUND(traffic2019.Traffic,2) AS 2019_traffic,
      ROUND(in2020.Traffic,2) AS 2020_traffic,
      ROUND((((ROUND(in2020.Traffic,2) -
ROUND(traffic2019.Traffic,2))/ROUND(traffic2019.Traffic,2))* 100),2) AS
`Percentage`
FROM (
      -- traffic in 2019
      SELECT pct.idMetropolitanCounty, mc.Name, pct.Traffic
      FROM tra0413.PedalCycleTraffic pct
      LEFT JOIN tra0413.MetropolitanCounty mc
      ON mc.idMetropolitanCounty = pct.idMetropolitanCounty
      WHERE pct.`Year` = 2019
) AS traffic2019
LEFT JOIN (
      -- Traffic in 2020
      SELECT pct.idMetropolitanCounty, pct.Traffic
      FROM tra0413.PedalCycleTraffic pct
      WHERE pct.`Year` = 2020
) AS in2020
ON in2020.idMetropolitanCounty = traffic2019.idMetropolitanCounty
ORDER BY `Percentage` DESC;
```

All questions were answered by the constructed queries, giving us light to understand the data collected from real life and giving us the possibility of using them for different purposes and applications.

Web application

A web application was built to present the results of the queries built based on our questions about the dataset. The application was built in NodeJS with Express framework and bootstrap for better styling and clear visualization of results.



Pedal Cycle Traffic Analysis

Data analysis in the tra0416 dataset made available in Road Traffic Transport by the Department of Transport

1. Which years had highest pedal cycle traffic

Year	Total Traffic
2020	27
2021	21.3
2022	20
2019	18.8
2014	18.8
2018	18.2
2017	18
2016	17.4
2015	17.4
2012	17.4

Pedal Cycle Traffic Analysis

Data analysis in the tra0416 dataset made available in Road Traffic Transport by the Department of Transport

1. Which years had highest pedal cycle traffic

2. Which regions had the highest pedal cycle traffic in last year?

Region	Total Traffic
Great Britain	6.3
England	5.6
North West	1.2
Yorkshire and The Humber	1.2
South East	1.1
London	1
East of England	0.8
West Midlands	0.7
South West	0.6

Pedal Cycle Traffic Analysis

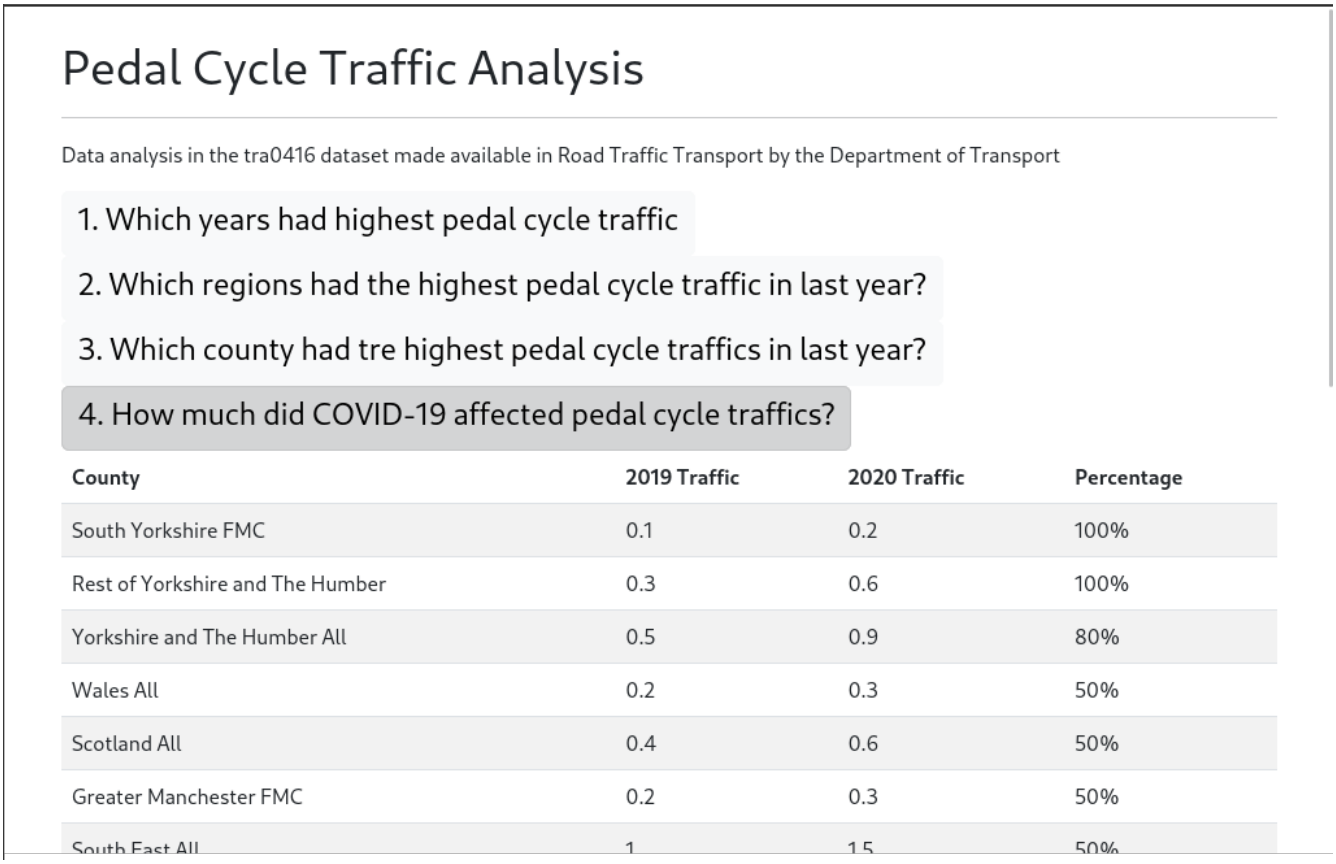
Data analysis in the tra0416 dataset made available in Road Traffic Transport by the Department of Transport

1. Which years had highest pedal cycle traffic

2. Which regions had the highest pedal cycle traffic in last year?

3. Which county had tre highest pedal cycle traffics in last year?

County	Total Traffic
Great Britain All	6.3
England All	5.6
South East All	1.1
London All	1
East of England All	0.8
North West All	0.6
South West All	0.6
Yorkshire and The Humber All	0.6



Pedal Cycle Traffic Analysis

Data analysis in the tra0416 dataset made available in Road Traffic Transport by the Department of Transport

1. Which years had highest pedal cycle traffic
2. Which regions had the highest pedal cycle traffic in last year?
3. Which county had tre highest pedal cycle traffics in last year?
4. How much did COVID-19 affected pedal cycle traffics?

County	2019 Traffic	2020 Traffic	Percentage
South Yorkshire FMC	0.1	0.2	100%
Rest of Yorkshire and The Humber	0.3	0.6	100%
Yorkshire and The Humber All	0.5	0.9	80%
Wales All	0.2	0.3	50%
Scotland All	0.4	0.6	50%
Greater Manchester FMC	0.2	0.3	50%
South East All	1	1.5	50%

Conclusion

Therefore, as shown in the results of the consultations, the year 2020 was very different from what was expected so far based on previous pedal traffic data. Traffic was much higher than expected. From this, we were able to find the regions and counties with the highest traffic, in addition to finding the percentage compared to the previous year. This data is shown in the developed web application and it can be important for understanding how the population reacted to the event and where there was the greatest impact. We can therefore make better decisions regarding urban planning and traffic in Great Britain.