

---

# Data-driven, risk-based road safety approach by sharing unique data on driving behaviour

*Menno Mimpfen\*, Thijs Jagt*

*ANWB, The Netherlands*

\* Corresponding author: [mmimpfen@anwb.nl](mailto:mmimpfen@anwb.nl) | <https://www.linkedin.com/in/mennomimpfen/>

## Abstract

ANWB introduced the Drive Safe ‘pay as you drive’ car insurance in 2016. Our goals: offer a car insurance that is much better aligned with the actual risk of the insured and improve driving behaviour, resulting in less accidents. Also support of road authorities to increase road safety has been our goal from the start.

In 2020 ANWB performed a pilot with City of Rotterdam, by providing aggregated and anonymous driver behaviour data, to improve accident risk calculations of their machine learning model. Because of successful results, a non-commercial data-service was launched in 2021 to support other road authorities in the same way. Over the last 6 months the number of road authorities having access to the data doubled to 160 in total. Data visualisation and machine learning models will increasingly support road authorities to determine where road infrastructure improvements are most efficient to reduce the number of road casualties.

## Keywords

road safety; innovation; driving behaviour; data visualisation; data analytics

## Road safety data: road authority’s needs

In many countries clear goals are set to reduce the number of traffic casualties. Next to the widely used Vision Zero for 2050, in The Netherlands a national road safety plan is in place to reduce the number of traffic accidents by 50% in 2030 compared with 2020. The need to use data to take well informed decisions will increase significantly over the coming years. This is due to the fact that the number of clear ‘black spots’ decreased. Black spots have been the main identifier to prioritise locations to enhance road design to improve traffic safety. These black spot locations are found by the total number of accidents over the past 3 years. Because the number of clear black spots is decreasing, additional tools are required to be able to take proper decisions to improve the design of the road infrastructure. High quality decision support information is required to determine where road infrastructure improvements will be most efficient to improve road safety. This requires the availability of appropriate data and tooling.

To move to a new way a working, the first challenge will be the availability of valuable data, including requirements such as:

- Data coverage for the full road network
- Homogenous data
- Availability over multiple years
- Data quality
- GDPR restrictions
- Affordability

---

## Road safety data: insurance company's needs

Driving behaviour has direct impact on the risk of being involved in an accident. As an insurer, we want to help our policyholders in changing their driving behaviour to reduce the risk of getting involved in a traffic accident. Measuring driving behaviour on an individual level, in terms of inappropriate driving speed, harsh braking, harsh cornering, severe accelerating and smartphone usage (while driving) enables us to assess the risk of getting involved in an accident.

### 'Pay how you drive' car insurance

#### Traditional car insurance

The premium for traditional car insurance is calculated based on a number of risk factors that affect the likelihood and extent of insured risks. Weight, for instance, affects the level of damage which can be caused by a car. So weight is a relevant factor in calculating the risk premium of the liability part of the car insurance. Another factor is car theft. The probability of theft (attractiveness and region) and the value of the car affect the premium for the theft part of the insurance. Most of these factors are static and therefore the same for all insureds within a segment.

However, as an insurer, we also know that the driving behaviour has a significant impact on the likelihood of being involved in a collision. With our traditional static factors we were not able to properly weigh this behavioural component. This changed with increasing technological capabilities for tracking and interpreting driving behaviour. Nowadays driving behaviour can easily be tracked and even influenced. This made it possible to make individual behaviour a more important risk factor in determining the level of premium.

#### How does 'pay as you drive' work behind the screens?

For our Drive Safe pay as you drive car insurance we collect driving behaviour data with a dongle connected to the car, or a mobile app which is linked with a beacon in the car. Based on the analysed data, we provide feedback to our insureds via a dashboard in their mobile app. This feedback is provided every 10 days and includes an overall driving score between 0 and 100. The higher the score, the safer the driving behaviour. It also includes textual feedback for each assessment factor as well as suggestions for improvement, if needed. The assessment factors are:

- Acceleration
- Braking
- Cornering
- Speeding
- Smartphone use (still in pilot phase)

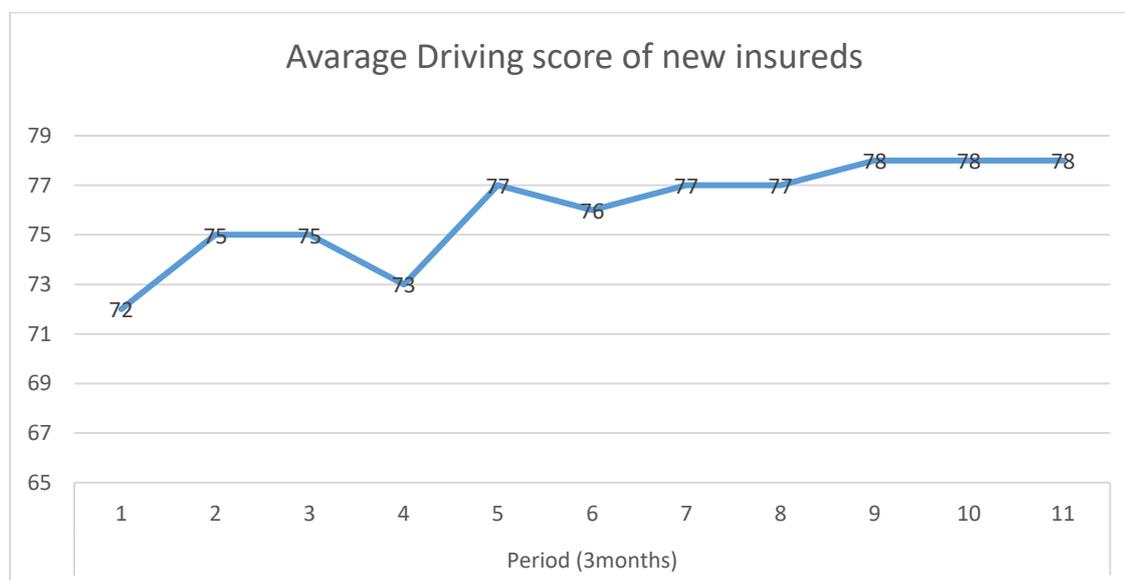
Every 3 months the overall driving score is used to calculate the discount on the insurance premium. By motivating our insureds to drive safe, the risk of being involved in an impactful traffic accident is reduced significantly. Because of this also the costs for the insurer should be lower compared to a traditional car insurance.

The 'pay as you drive' car insurance exists for 6 years now. Overall, we have seen the cost of claims decreasing significantly. This cost reduction is equal to the financing investments required. These include the development and maintenance of used technology as well as the reduced income because of the discounts on the insurance premium, which can amount up to 30%.

## Results

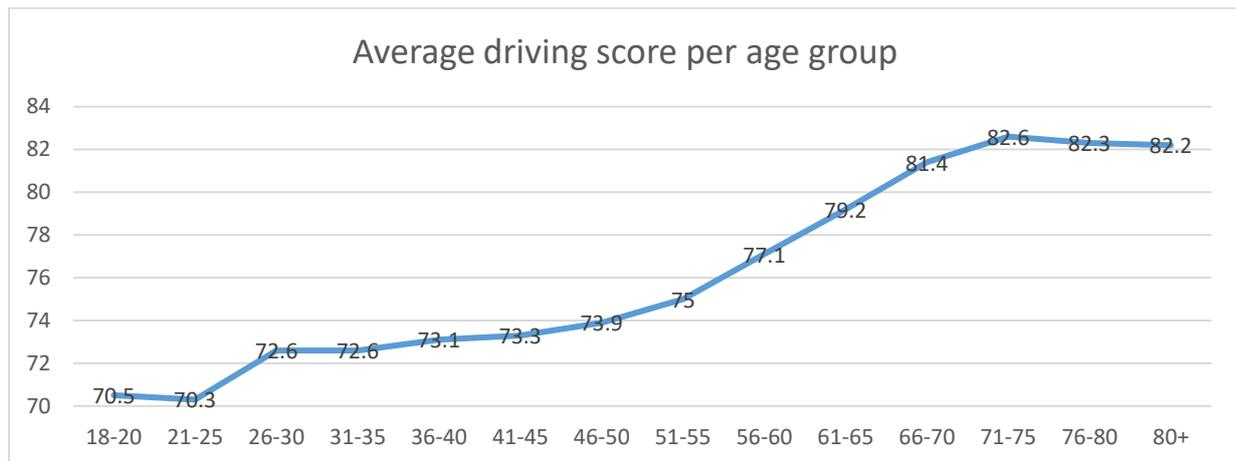
We have been able to identify patterns in our customers' driving behaviour. At the start, new Drive Safe insured's driving behaviour improves. However, after a while, many of our customers get used to the system and our feedback. It loses their attention, and their driving score decreases. As a result, the discount decreases as well, and customers become aware of their negative change in driving behaviour. Next, they recover to the initial improved level and improve even further. This pattern confirms the positive effect of encouraging our insureds to improve their driving behaviour through continuous feedback and a premium discount.

The graph below presents the average driving score per quarter of new insureds, and shows the pattern mentioned above. Their average score after the first quarter is 72 on a scale between 0 and 100. After an improvement in the second and third quarter the driving score decreased and finally improves again to score around 78.



Some additional results and insights based on the Drive Safe data:

- 95% of our customers is rewarded by a discount
- The average driving score is 76, which corresponds to a 16,6% discount.
- Kia Niro drivers have the highest average driving score (83.0) while drivers of a BMW 3-series have the lowest (69.9).
- The average driving score between the age of 71 to 75 years is the highest (82.6%). See graph below for full details.
- Over a period of six years, we have collected more than 2 billion km of driving behaviour data.



In addition to the feedback we provide to all our customers via the mobile app, we are piloting additional coaching for those with the lowest driving score. We are supporting them with additional (one-on-one) coaching to raise awareness of the increased risk and help them to improve driving behaviour. So far, the pilot has been successful. The response of the approached customers was positive.

In conclusion, road safety of individual drivers is improved because of our Drive Safe car insurance, and the insurance premium is much more aligned with the actual risk.

### Closing the gap by public private partnership

From the start it has been ANWB’s vision to use the data of the Drive Safe car insurance to improve road safety in general as well. Because of this vision, the use case was covered from the start in the Terms and Conditions of the insurance policy. Potential customers as well as insureds have been pro-actively informed about our plan to share the data in an anonymous form with road authorities to contribute to improvement of road safety.

Aggregated car driver behaviour data is very helpful to analyse and identify safety issues in the road network. For example, frequent harsh braking on a specific road section can be the consequence of the design of the road infrastructure, resulting in traffic accidents. Since our specific group of car drivers wants to drive in a safe way this data is even of more value.

In 2019 the municipality of Rotterdam proactively looked for data to improve their machine learning model which determines the accident risk for each road section in the city. This resulted in a first use case for the aggregated data.

City of Rotterdam mentioned: *‘If even the safest drivers still make a massive mistake somewhere, then there may be something wrong with the road infrastructure’.*

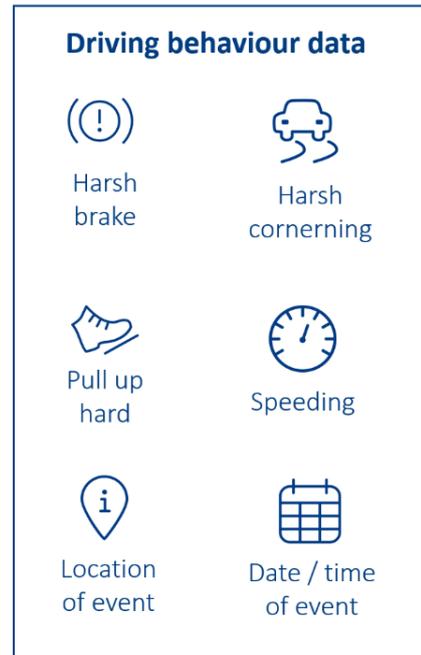
Since majority of the costs for the developments and technology to provide the data is covered by our insurance product, ANWB is willing to offer data as a non-commercial service to road authorities under the condition it is being used to improve road safety. Our societal challenge resulted in multiple innovations already, with Public Private Partnership as key for success: floating car data from the car insurance is provided via a non-commercial data-service to road authorities.

## ANWB's data service for road authorities

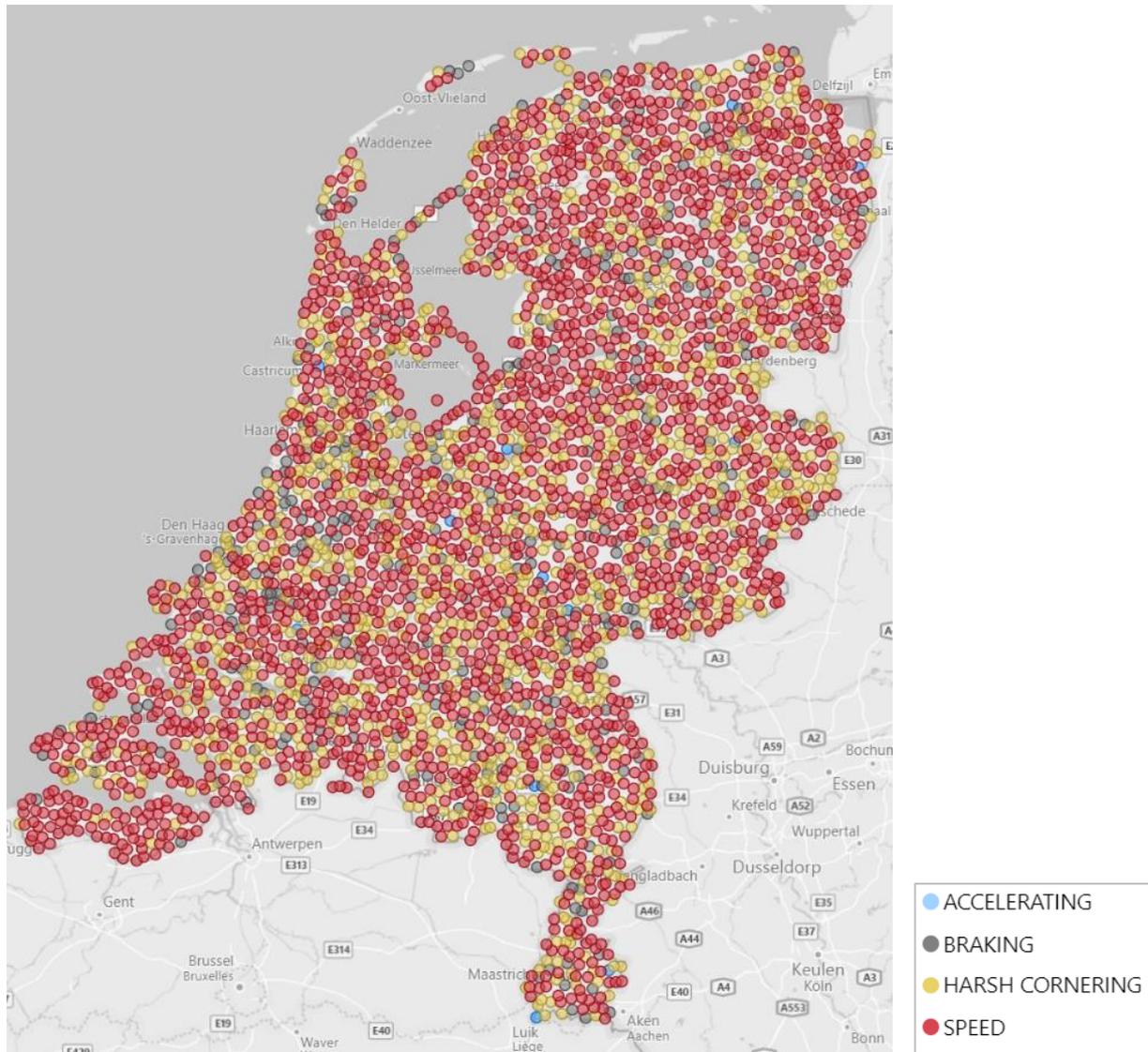
ANWB's ambition is to move to zero road fatalities in 2050 and contribute in multiple ways. By providing valuable data from our Drive Safe car insurance, new insights become available for traffic consultants of road authorities. By moving to data driven decision making, road authorities also contribute to the national ambition of governments in The Netherlands to change to a risk-driven approach to improve road safety.

The raw data like GPS location and g-forces from more than 75.000 cars is translated into useful information to calculate the score of each individual driver. Aggregated and anonymous data including all events has been archived from the start in 2016. The image shows a quick overview on the data being provided to road authorities. We only share these events, meaning if all insureds would drive perfectly, we wouldn't have any data to share with road authorities! For each event also include the severity and more addition detailed information like maximum g-force and speed during the event.

Since early 2021 the data is available for road authorities via an easy to use and secured data-service. A download link to an encrypted ZIP file is sent by email to our contact at a road authority. At the moment an encryption key is sent to our contacts mobile phone, to get access to a ZIP file. Every month, after a quality check is done, a new dataset is available.



Driver behaviour events are only stored at the moment something unusual happened during a trip. The map below shows a reduced set of the locations where those events happened over the first quarter of 2021. In fact, there was much less traffic during this period because of the COVID pandemic. In case our insureds could drive 100% perfectly, this map would have been empty.



*Nationwide driving behaviour events during first quarter 2021*

### Use cases

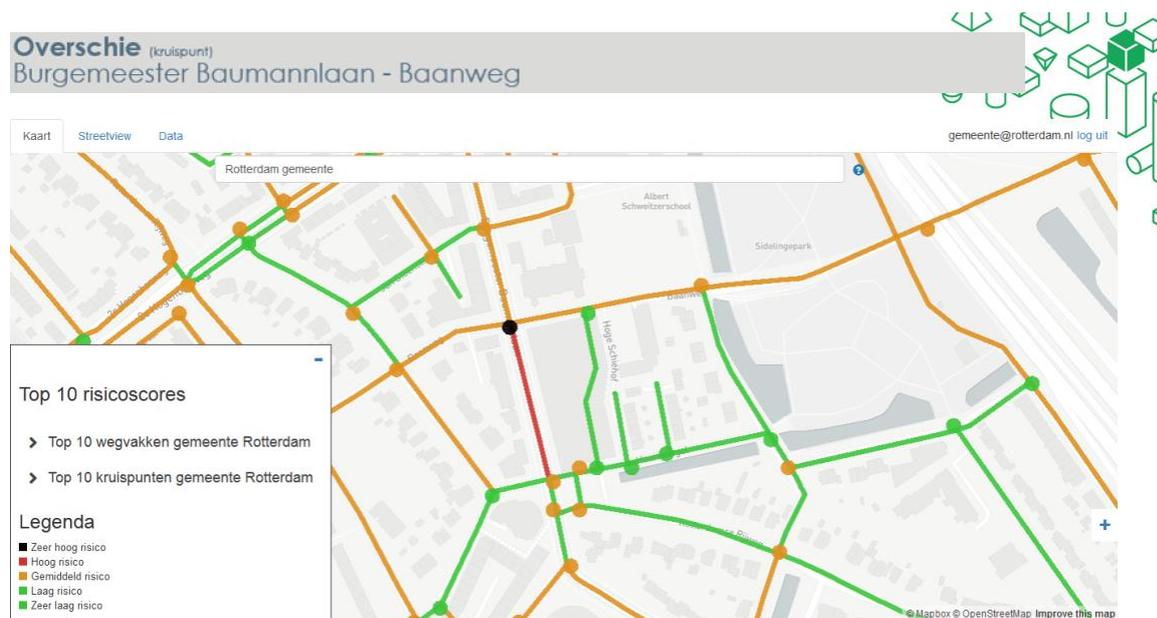
Adequate road design helps to prevent human mistakes in traffic, and less human errors will result in less accidents. It's important of course to use the limited budget of road authorities to improve road design as effective as possible. Meaning well-supported decisions making is required to prioritise improvements in road infrastructure.

Black spots have been the main identifier to determine the priority of locations to change road design to improve traffic safety. These black spot locations are identified by the total number of accidents over the past 3 years. Because the number of clear black spot locations is decreasing additional tools are required to be able to take proper decisions while prioritising improvements in road infrastructure.

### Using machine learning to determine accident risk

In 2019 the municipality of Rotterdam initiated an innovation to make use of machine learning to calculate accident risks. This was done for each road section and crossing in Rotterdam, see the map showing an example below. Colour codes are used to present the accident risk. Black and red indicate high risk, orange average risk and green stands for a low risk.

In 2020 a cooperation started between ANWB and Rotterdam to add driver behaviour data to the machine learning model. Most important learning from this pilot was that driving behaviour data is an important predictor for traffic accidents. Especially for locations with a high number of harsh braking events.



Information from machine learning models is being used by a fast growing number of road authorities in The Netherlands. Currently the output of the model is used to improve prioritisation of the black spot locations. Developments are ongoing to further improve the information from the models as well as how to make best use of the outcomes.

### Big data visualisation

Next to making use of data for machine learning, data visualisation is an increasingly important tool to make best use of the terabytes of data being generated. A good visualisation removes the noise from data and highlights useful information. Even basic data visualisation helps already, it enables the further discovery of value which is still hidden and even new use cases can be discovered by visualising and filtering the data in a smart way. At ANWB we started a new initiative to support road authorities in their basic need to visualize driving behaviour data.

Even without any additional data science or smart filters, the driving behaviour data already provides valuable insights. For example, the roundabout below shows a mixture of frequent driver behaviour events. A quick look at this location probably explains the cause: tall bushes in the middle of the roundabout.



*Local snap shot from nationwide driving behaviour showing multiple event clusters*

---

Use cases we identified so far related to visualisation of the driver behaviour data include:

- Better insights whether an accident location is related to common driver behaviour, what type of behaviour is dominant and where this behaviour exactly occurs.
- Monitoring drivers' behaviour as an effect of changes in road infrastructure. After improvements in the infrastructure, adjustments in regulations and/or enforcement, monitoring of changes in driving behaviour is in place immediately, without any additional effort.
- Verification and better understanding of citizens' complaints about road safety based on data on actual driving behaviour.
- Decision support in the process to determine next most effective improvements in road infrastructure.
- Enabling discovery of new use cases while using road safety related data by experimenting and combining multiple data sources.

## Results

Starting with one user of the data-service early 2021, and 76 by the end of 2021, the number of road authorities doubled in 2022 already to a total of 160. Many road authorities started to include the data and risk driven approach into their decision process of prioritisation as well. Next to this, a large number of road authorities is still in the process of change and development.

Big data visualisation and machine learning models will become more mature. They will provide high quality decision support information to determine where road infrastructure improvements will be most efficient to improve road safety as well as the measures to be taken to improve road safety.

## Developments

The number of road authorities having access to ANWB's driving behaviour data grew rapidly over the past 1,5 years and reached the total number 160 in August 2022. This also results in a growing feedback and ideas how to further develop value and extend the service.

Data developments ANWB will be focussing on in 2022-2023 include:

- Extension of the data regarding usage of a mobile phone while driving to get a better understanding of distraction by mobile phone usage. This includes the locations where car drivers use their phone to make calls (hands free or not is being distinguished) or touch the screen of their phone.
- Research and development concerning usage of floating car data to determine hot spot locations for turn arounds on provincial roads and take measures to improve road safety.
- Research and development of floating (e)-bike data, focussing on insights in driving behaviour of cyclists and traffic flows to improve road safety.

## References

Soleymanian M., Weinberg C, Zhu T. (2017) *Sensor Data, Privacy, and Behavioral Tracking: Does Usage-Based Auto Insurance Benefit Drivers?*

[https://news.ubc.ca/wp-content/uploads/2017/06/UBI\\_Paper\\_Latex\\_Marketing\\_Science-with-name.pdf](https://news.ubc.ca/wp-content/uploads/2017/06/UBI_Paper_Latex_Marketing_Science-with-name.pdf)