Challenges cloudera:

Challenge1: Volume/Menge an Daten:

Per car:

- 2 petabytes of data per year total
- Significant data:
  - 25 GB per hour
  - 130 TB per year
- Flexible storage store any and all date in Kudu and HDPS
- Data used for data warehouses and real time applications

Acquisition and analysis of more than 30000 signals and data points from sensors

Challenge2: Variaty

Various data types:

- Data from ECU
  - Speed, fuel, temp, brakes etc.
- Location
- Safety data
- Camera recordings (computationally intensive)

Data must be partially streamed in real time.

- Data Sources:
  - ECU Electronic control Unit
  - Vehicle Plug ins
  - Head units
  - Cameras

Challenge3: Velocity

Separation into On Edge and Cloud Analytics

- Edge: Fast immediate calculation necessary, braking times, accident detection.
- Cloud. Computationally-intesive analytics, mashine learning, time series, trends.

Challenge4: Veracity:

Four pillars of security. perimeter, access, visibility, data and recording service.

Challenge5: Value

the most important use cases:

- Predictive maintenance
- Usage-based insurance
- Public services

Targets:

- 150 trillion in revenue
- 250 million connected vehicles
- Quadruple revenue

Benefits:

- 80% fewer alcohol-related accidents
- Predict maintenance intervals
- Automatic driving
- Entertaiment
- Comfort and safety

Challenges level 1 to level 4

Level 1: Data Source Layer

Very many data sources both structured and unstructured.ECU

- Vehicle Plug ins
- Head units
- Cameras
- Problem: Many different data types

Tools:

- SQL
- NoSQL
- Kafka
- Flume..
-

Level 2: Data Storage Layer

Hadoop Hbase (NoSQL) storage concept:

- Streaming vehicle data
- Geolocation
- Manufacturing Supplier
- Parts and warranties
- Maintenance data
- Dealer data
- Customer data

All this data is collected for later analysis.

Problem: New data must be read in every 5 seconds.

Level 3: Processing Layer

- Direct access to Data Storage:
  - Statistical methods
  - Machine Learning
  - Artificial Intelligence and Deep Learning

Potential use cases:

- Prdictive maintenance to improve performance and reduce downtime for fleets.
- Based on insurance to reduce claims by a major European insurance agency
- And many more

Level 4: Data Output Layer

-
- Visualization via cloudare BI website.
    - Calculation of aggression value to compare different drivers.
    - Comparison accidents vie aggression values and correlation between this value and oil/brake replacement.
    - And much more.