

EPIC DATA CONSULTING

20-25 Minuten Pitch

Unser Team für Sie

Nicolas Mahn, 263020



Jan-Hendrik Hausner, 261056



Artur Erfurt, 262311



Kevin Hilzinger, 263023



Patrick Jäckle, 263032



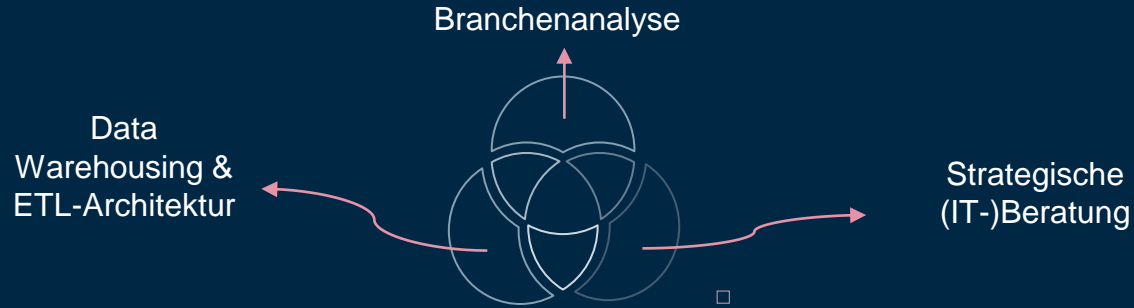
Niko Kauz, 263616



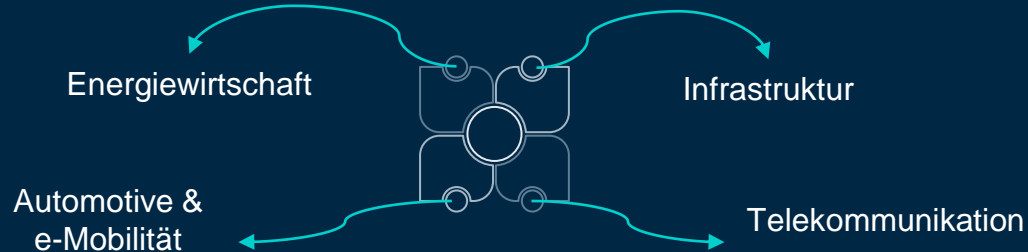
**“Das Ziel ist es, Daten in
Informationen und
Informationen in Erkenntnisse
zu verwandeln.“**

— CARLY FIORINA, EHEMALIGE CEO, HEWLETT PACKARD

Kompetenzfelder



Branchenwissen



Unsere Epic Data im Profil



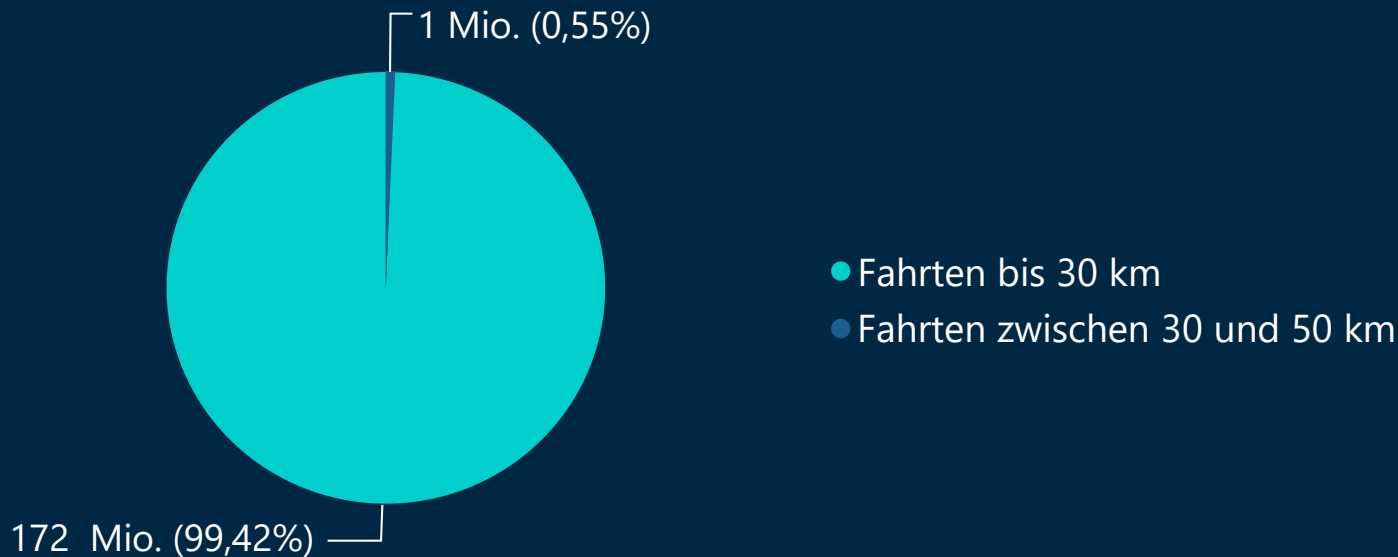
Unsere umfangreichen Daten-Assets ermöglichen uns Sichten auf Ihre Branche, welche der Konkurrenz verborgen bleiben

Was die Daten in der Analyse sagen:

- Größter Teil der Fahrten unter 30 km
- Ladestationen und Flottenpark sollte einen Fokus auf den Norden haben
- Erfassungsfehler der Daten bei manchen Verifone-Transaktionen
- Flottenzusammensetzung mit Priorität auf Kleinstwagen
- Zahlungsausfall in Kalkulation berücksichtigen



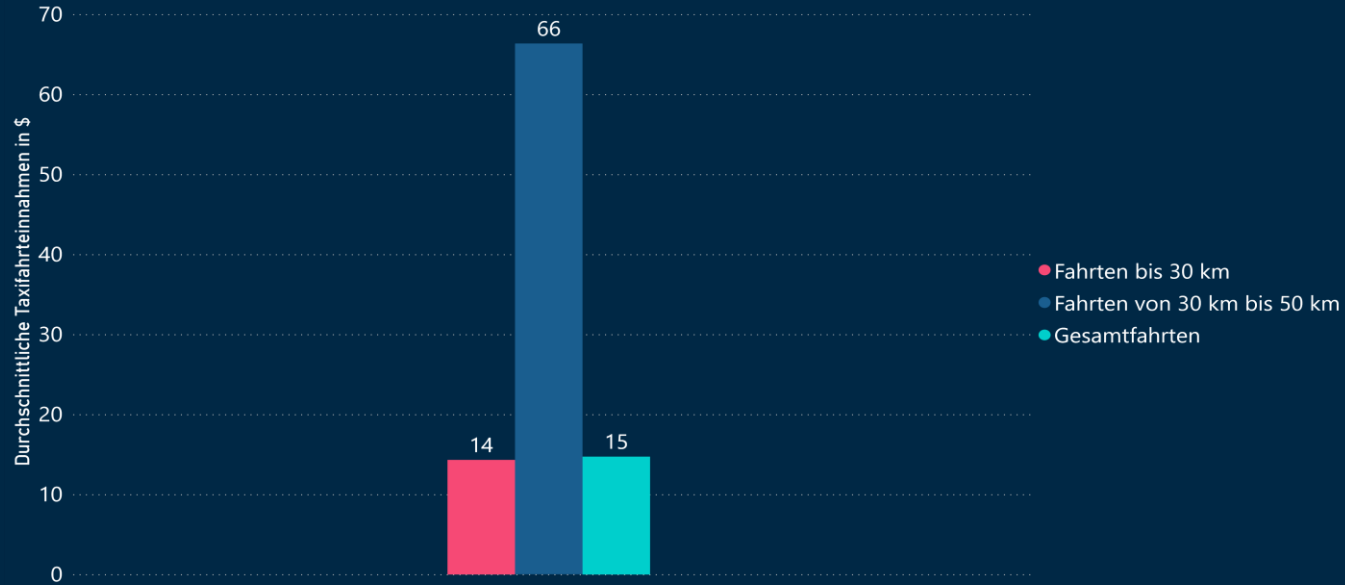
Taxifahrten insgesamt



- Mehrheit der Fahrten unter 30 km
- Anschaffung der E-Taxis mit 30 km Reichweite



Verdienst pro Fahrt

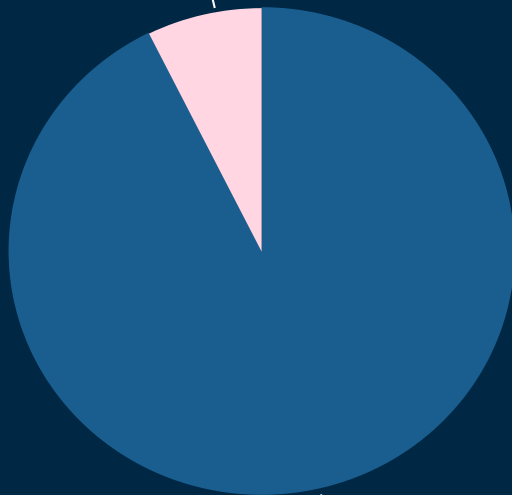


Fahrten zw. 30 km und 50 km haben den höheren Verdienst



Fokus auf den Norden

13 Mio. (7,34%)



160 Mio. (92,66%)

- Startpunkte im Norden
- Startpunkte im Süden

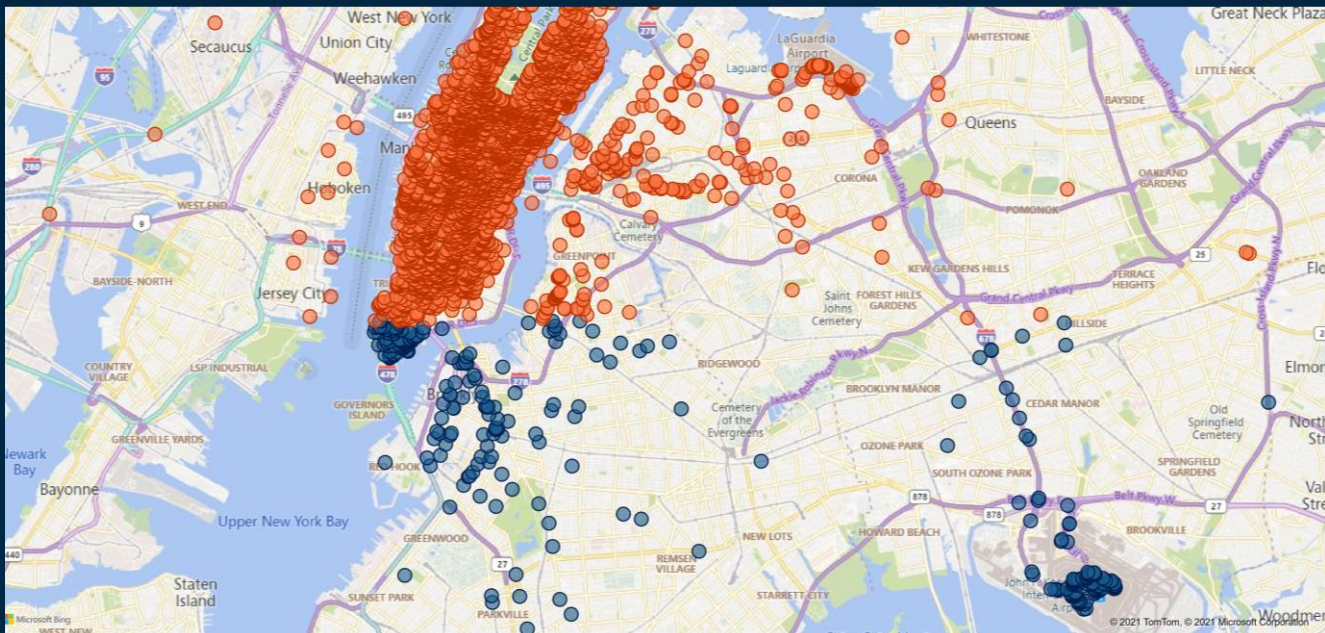


Ladestationen für E-Taxis im Norden positionieren





Die Fahrten in NYC veranschaulicht



Wahl des Flottenparks mit Schwerpunkt im Norden

► „Das Nächste nehme ich“

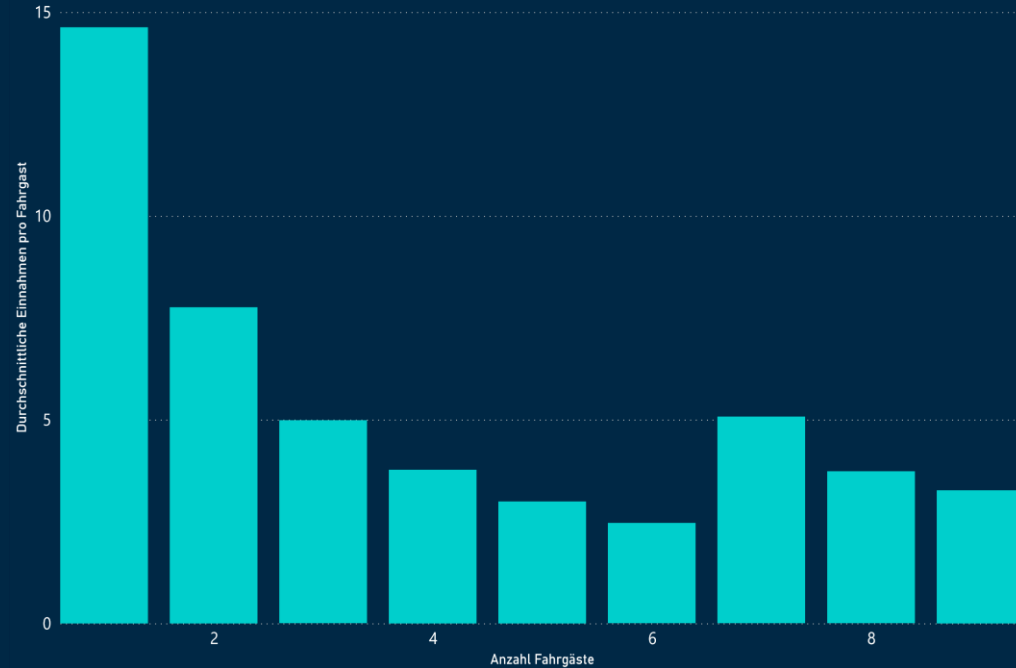


Im Durchschnitt wird eine Cab-Fahrt nur von 1,71 Personen auf einmal in Anspruch genommen.

- • Künftige Flottenzusammensetzung mit hohem Kleinwagen-Anteil
- PaaS-Portal zum Cab-Sharing in Erwägung ziehen

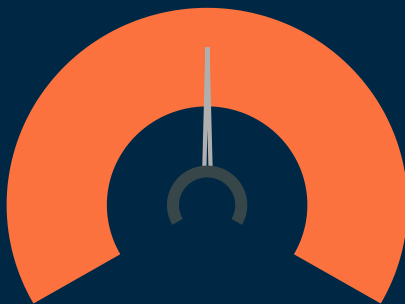
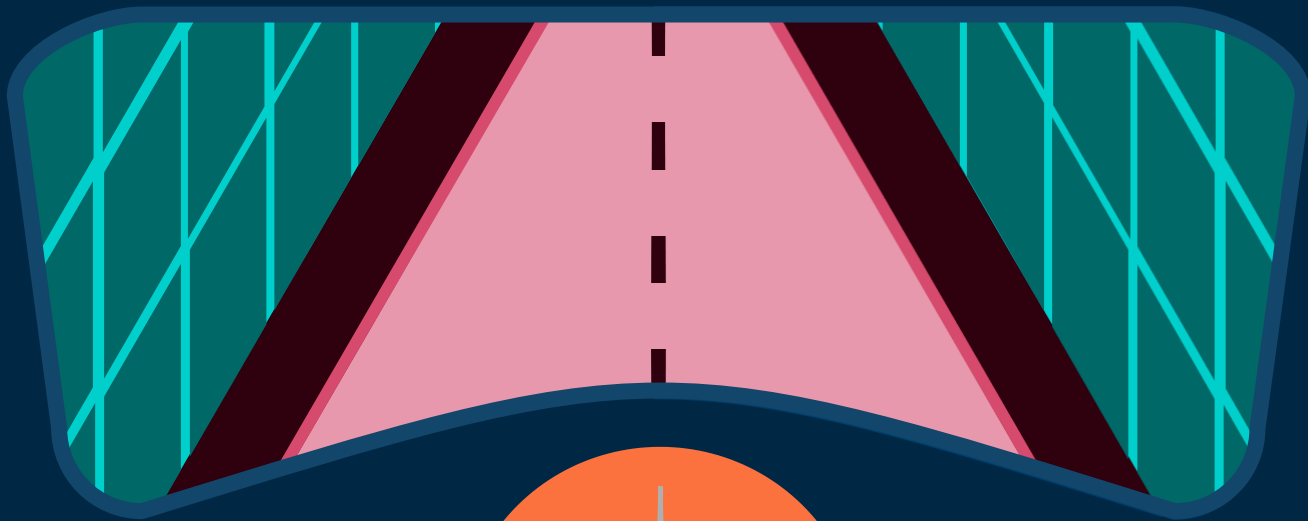


Einnahmen pro Fahrgast



Auch größere E-Taxis sollten beschafft werden

Durchschnittsgeschwindigkeit

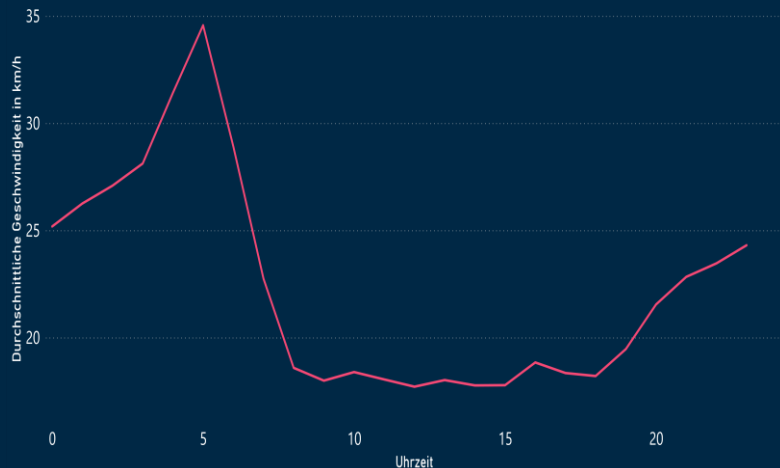


20,89 km/h



Korrelation zwischen Geschwindigkeit und Verkehr

Geschwindigkeit



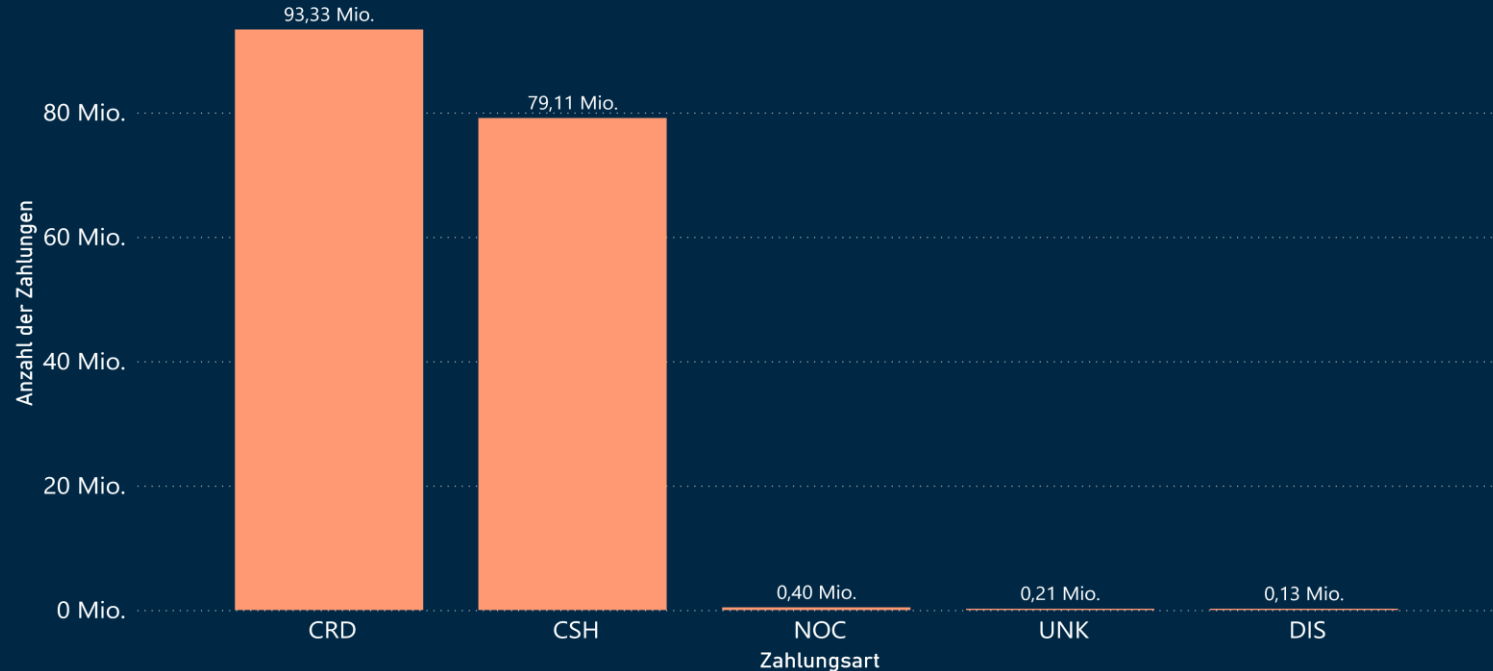
Verkehr



- Geschwindigkeit ist in Relation zum Verkehrsvolumen
- Tarife sollten auf die Tageszeit angepasst werden



Nutzung der Zahlungsmethoden



Ein Zahlungsausfall von 0,31% sollte in die Tarife mit einberechnet werden



Erfassungsfehler der Daten bei VTS

Verifone[®]

86.497.730 Ausgestattete Fahrten insgesamt
206.874 Fahrten mit unbekanntem Transaktionstyp

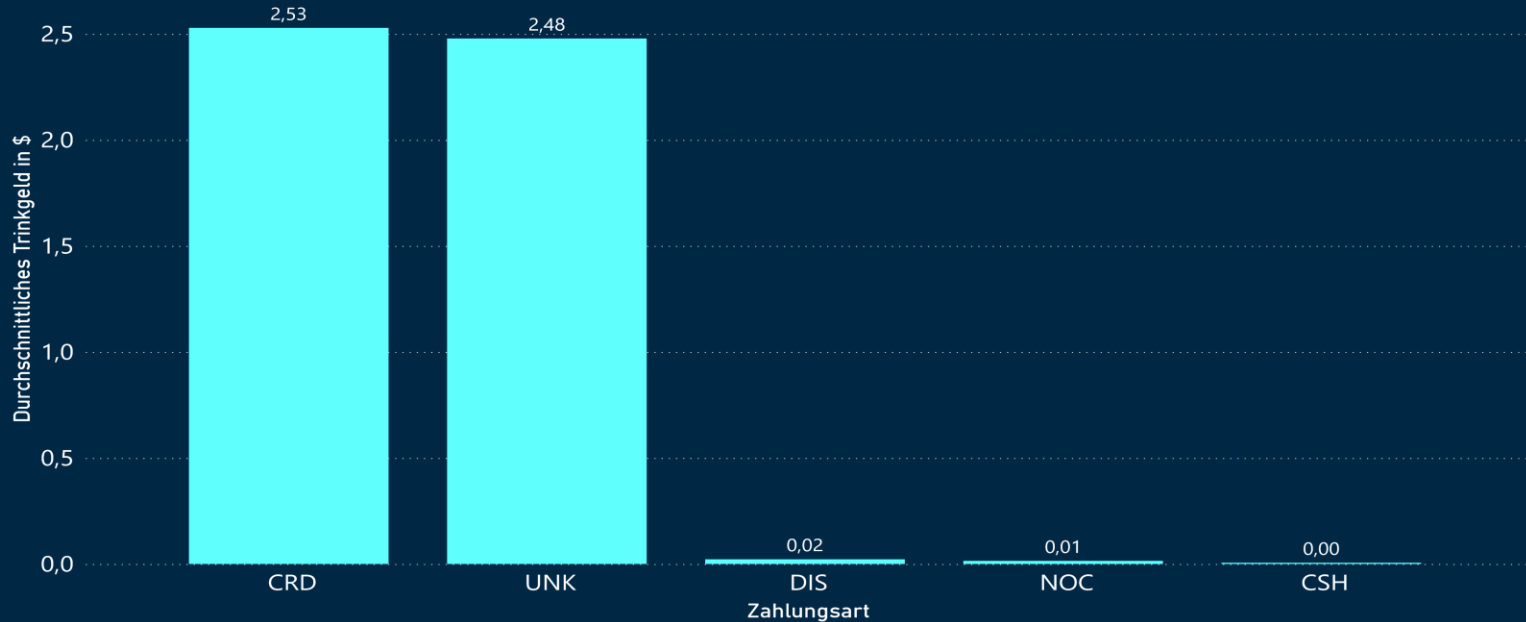


86.684.657 Ausgestattete Fahrten insgesamt
0 Fahrten mit unbekanntem Transaktionstyp





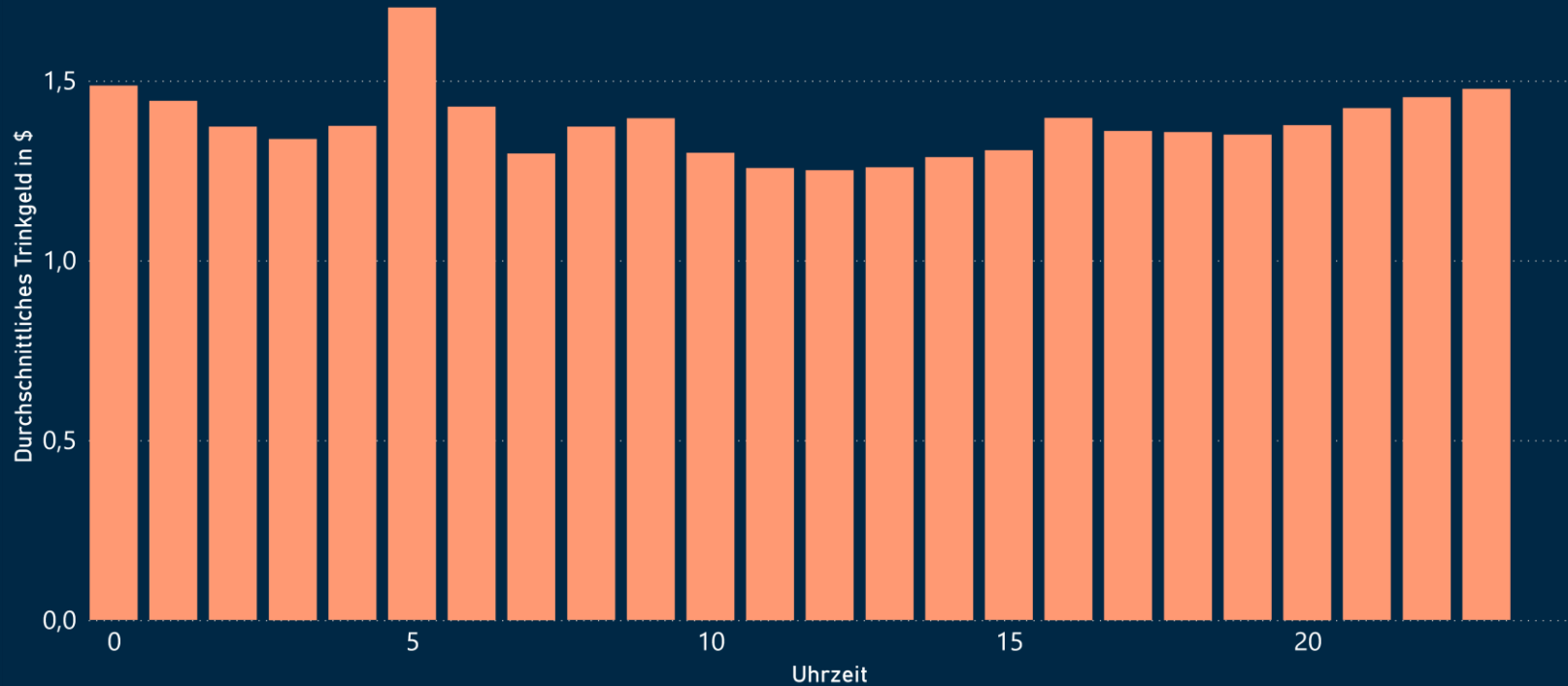
Durchschnittliches Trinkgeld nach Zahlungsart



Maßnahmen zur Erreichung eindeutiger Zahlungs-
zuordnungen einleiten

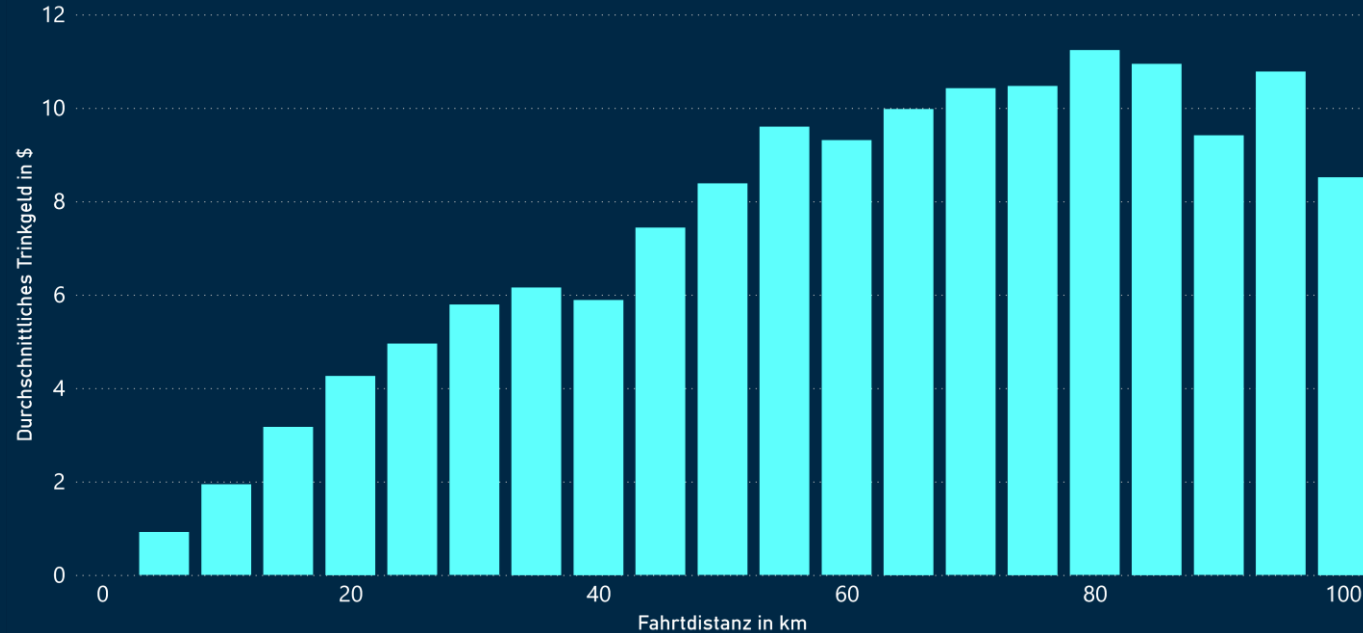


Trinkgeld über den Tag





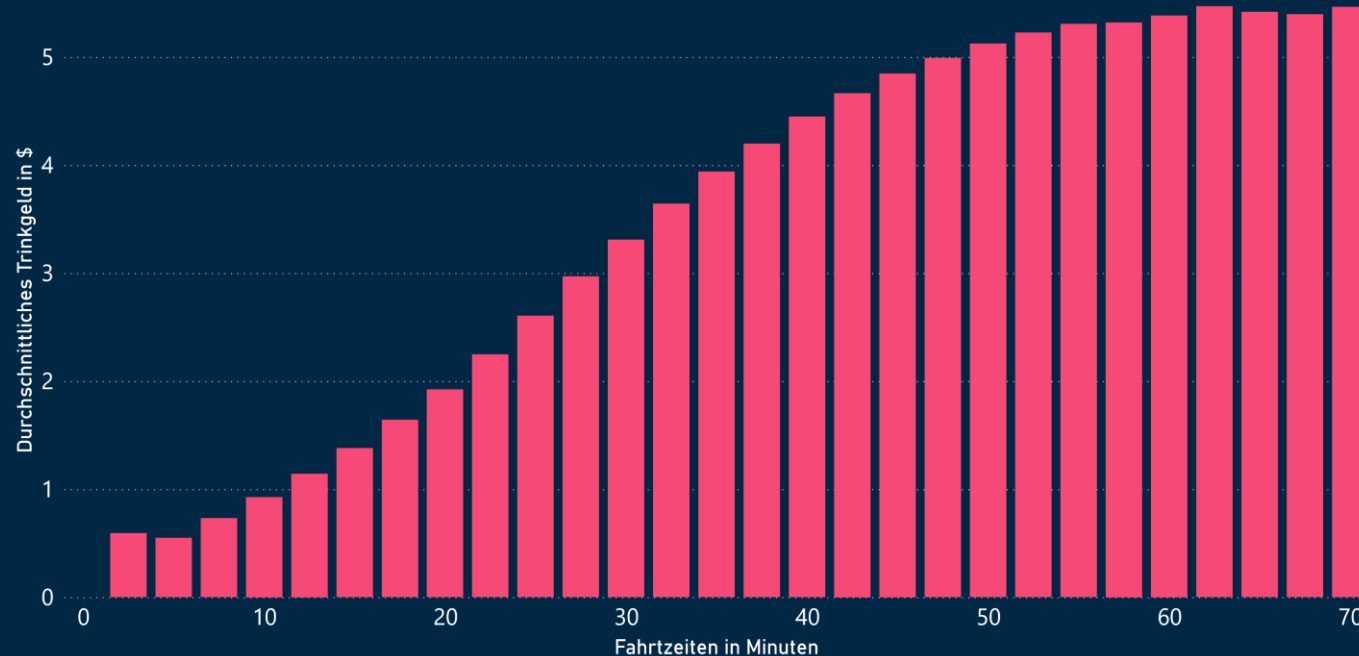
Mehr Trinkgeld bei längerer Fahrdistanz



Trinkgeld steigt mit der Fahrdistanz, nimmt aber bei hohen Fahrdistanzen leicht ab



Mehr Trinkgeld bei längerer Fahrtdauer

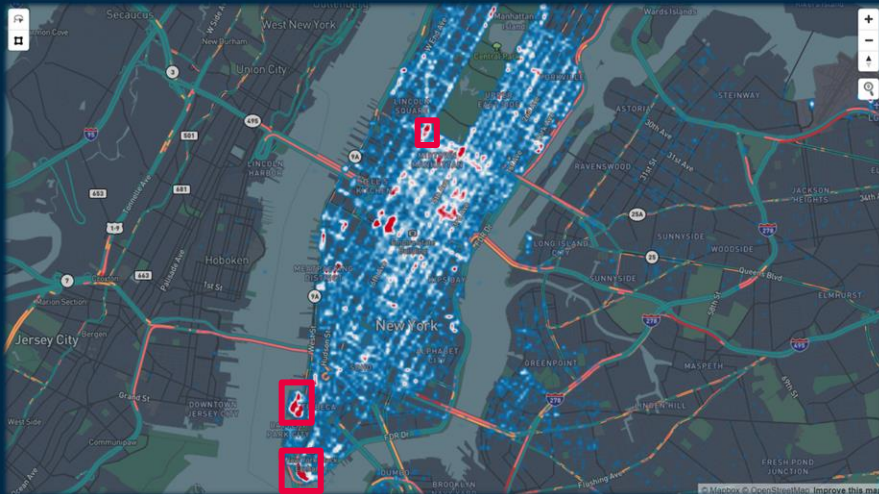


Trinkgeld stagniert ab einer Fahrtdauer von 50 Min.



Orte mit unbekannten Zahlungsvorgängen

Aufkommen an Fahrten mit unbekannter Zahlung



Geographisch bedingte Verkehrsdichte



- Häufung von Fahrten mit unbekannten Zahlungsvorgängen
- Spezielle Betrachtung von Punkten, die sich nicht mit der allgemeinen Verkehrsdichte decken

Technologien

Apache Spark



Einheitliche Analyse-Engine für die Verarbeitung umfangreicher Daten

- 100x Schneller als Hadoop MapReduce
- Batch Applikationen und Berechnungen, iterative Algorithmen, interaktive Queries und Streaming in einer Engine
- Verarbeitet Daten in Resilient Distributed Datasets (RDDs), DataFrames und Datasets
- Organisiert die Verteilung der Aufgaben/Daten für Berechnungen in einem Cluster
- Leichte Einbindung über APIs (Spark SQL für Java)

Mit APIs zur Einbindung, Orchestrierung der Nodes in einem Cluster und Batch Funktionen erleichtert Spark die Skalierbarkeit von Berechnungen und Auswertung von enormen Datenmengen

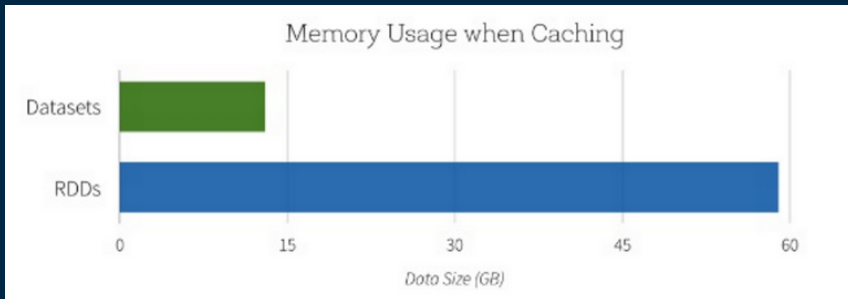
➔ Starkes Tool für Big Data



DataFrames und Datasets

... das Upgrade zu RDDs

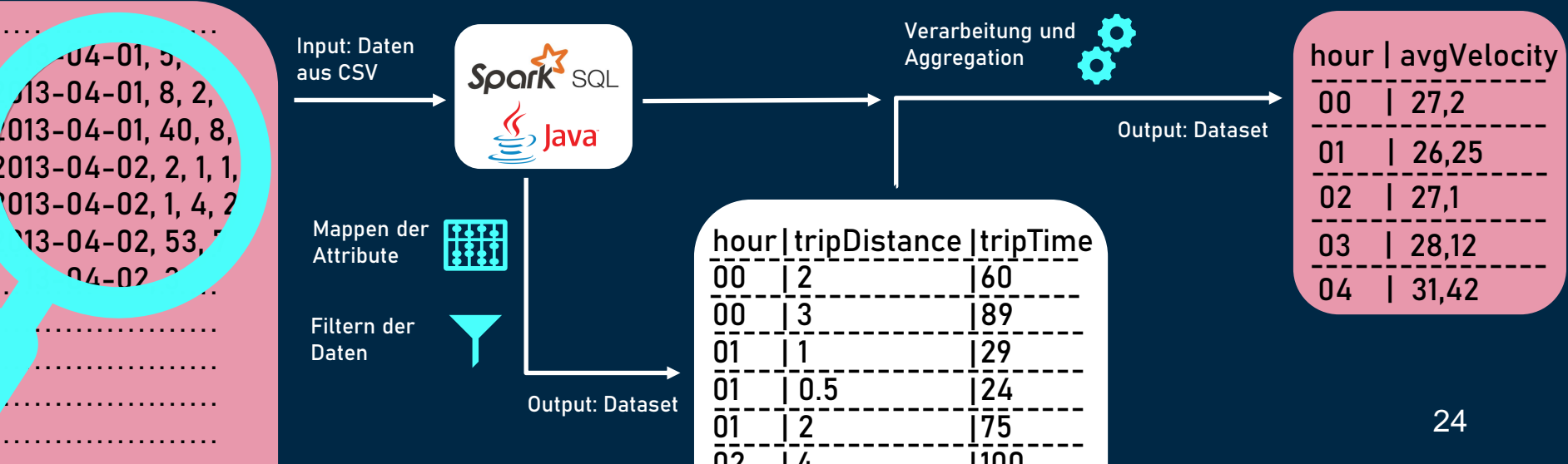
- Nutzung der optimierten Spark **SQL Execution-Engine**
- **Unveränderliche, verteilte Sammlungen** von Daten
- **Strukturierung** der Daten in Spalten mit **festen Datentypen**
- Designed um noch **größere Datenmengen** noch **schneller zu verarbeiten** mit komplexeren Möglichkeiten
- Erkennung und Behandlung von **Fehlern**
- **Effizienter** im Platzverbrauch



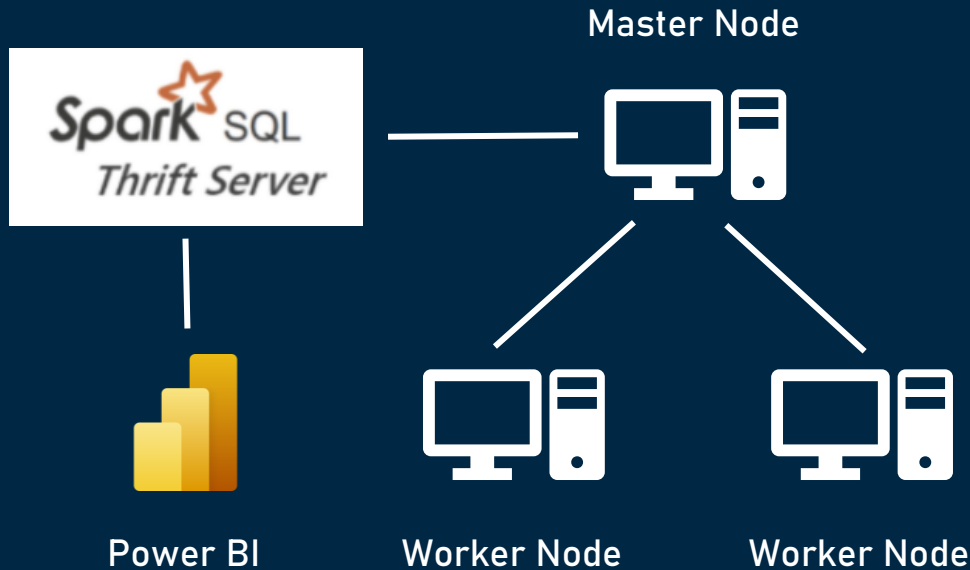
Spark SQL

- APIs zum Ausführen von **SQL Queries** und **Samplen** von Daten
- Verarbeitung und Ausgabe der Daten in **DataFrames** und **Datasets**
- Liefert mehr Informationen über die **Strukturen** und **performiert** damit **optimierter**
- Interaktion und Anbindung über JDBC/ODBC für **Java** und **BI-Tools**

ETL-Strecke



Architektur



Spark Master at spark://hennih... x

Not secure | hennih...ddns.net:8080

Spark Master at spark://hennih...fritz.box:7070

URL: spark://hennih...fritz.box:7070
Alive Workers: 2
Cores in use: 16 Total, 16 Used
Memory in use: 28.7 GiB Total, 4.0 GiB Used
Resources in use:
Applications: 2 Running, 0 Completed
Drivers: 1 Running, 0 Completed
Status: ALIVE

Workers (2)

Worker Id	Address	State	Cores	Memory	Resources
worker-20211122075709-172.18.143.84-44731	172.18.143.84:44731	ALIVE	8 (8 Used)	14.3 GiB (1024.0 MiB Used)	
worker-20211122085630-192.168.188.21-45129	192.168.188.21:45129	ALIVE	8 (8 Used)	14.4 GiB (3.0 GiB Used)	

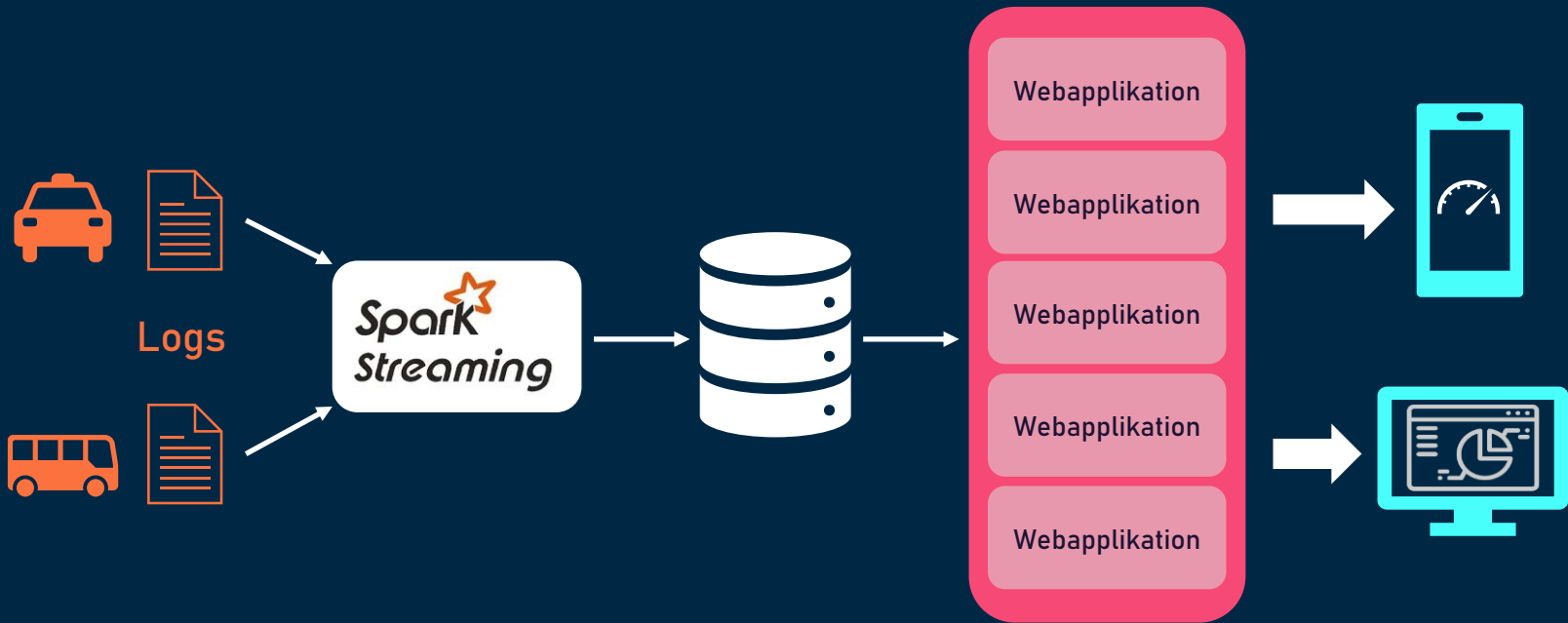
Running Applications (2)

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
app-20211122085824-0001	BusinessIntelligence	13	1024.0 MiB		2021/11/22 08:58:24	hennih...	RUNNING	1.8 min
app-20211122085650-0000	Thrift JDBC/ODBC Server	2	1024.0 MiB		2021/11/22 08:56:50	hennih...	RUNNING	3.4 min

Running Drivers (1)

Submission ID	Submitted Time	Worker	State	Cores	Memory	Resources	Main Class	Duration
driver-20211122085821-0000 (kill)	2021/11/22 08:58:21	worker-20211122085630-192.168.188.21-45129	RUNNING	1	1024.0 MiB		de.hfu.businessintelligence.Application	1.8 min

► Architekturempfehlung



➡ Daten in Echtzeit verarbeiten



Ziele gemeinsam erreichen



Konzeption und Implementierung PaaS-Plattform für Cab-Sharing



Rebranding – Ausrichtung der CI auf Nachhaltigkeit und Green NYC



Unternehmensinterne Datenanalyse zur Potentialerkennung



Vergleich der Unternehmens- und Branchendaten als Benchmarking



Umsetzung der neuen DWH- und ETL-Architektur



Beratung rund um staatliche Förderung der Infrastruktur und Flotte

Ihr Kontakt zum Erfolg



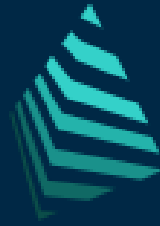
Niko Kauz
Senior Manager – Data & Integration

T +49 177 60358817

M +49 800 1111111

@ niko.kauz@epic-data.com

Zögern Sie nicht, uns zeitnah zu kontaktieren. Gerne besprechen wir in einem weiteren Termin die Details unserer Partnerschaft.



EPIC DATA CONSULTING

Backup Slides

Unser Code zum Einsehen



<https://github.com/1Zero64/BusinessIntelligence>



Erfassungsfehler der Daten bei VTS

- TPEP = Taxicab Technology Service Provider
- CMT = Creative Mobile Technologies LLC
- VTS = Verifone Transportation Systems von Verifone

TPEP	Payment Type: Unknown	Payment Type Known
CMT	0	86684657
VTS	206874	86290856



Softwaresysteme ausschließlich bei Verifone verbessern

Exkurs Cyber Forensic



Differenz zwischen durchschnittlichen
Einnahmen des Zahlungstyp „unknown“
und anderen Typen

Payment Type

3,66 \$

Other Types

Payment Type

Unknown: 17,78 \$

Exkurs Cyber Forensic

Dringende Empfehlung:
Vermeidung unbekannter Zahlungshistorien

- Prüfung auf Korrelation zw. Cab und Frequenz unbekannter Zahlvorgänge
- Regelintervall der Wartung der Zahlssysteme



Payment Type

Unknown: 17,78 \$

Beratungskategorien



Daten auswerten



Potenzielle
Wettbewerbsvorteile erkennen



Empfehlungen treffen

Kompetenzfelder



Data Warehousing &
ETL-Architektur

Branchenanalyse



Strategische
(IT-)Beratung



Evtl. auch Beratung bzgl. Förderung

Branchenwissen

Energiewirtschaft

Automotive & eMobilität

Infrastruktur

Telekommunikation