

- bachelor
- Ich habe mein ↑ Studium im Jahr 2020 in Pakistan abgeschlossen
  - Danach habe ich etwa einenhalb Jahr als Associate Data Analyst bei Systems Limited gearbeitet
  - Während dieser Zeit habe ich sechmonatiges Data Analytics - Training absolviert completed und danach habe ich Power BI verwendet used um Bericht und dashboard zu erstellen
  - Meine Aufgaben waren Daten analysieren KPIs darstellen represent und klar interactive Reports erstellen

- Im Jahr 2023 bin ich für mein Masterstudium nach Deutschland gekommen
- Während meines Studiums habe ich auch deutsche Sprache gelernt
- Außerdem habe ich als Werkstudent im Bereich DataAnalytics bei Siemens Healthineers gearbeitet
- Dort habe ich hauptsächlich mit Power BI Reports und dashboard gearbeitet
- Mein Schwerpunkt <sup>focus</sup> lag auf der Entwicklung und Optimierung der Performance von Report.
- Bei Siemens habe ich im Kunden service gearbeitet - Ich habe die Daten über die Lizenzen von medizinischen <sup>equipment</sup> Geräten gesammelt und einfache Dashboards erstellt
- Durch diese Dashboards konnte das Team einfach sehen, welche Land welche Geräte - Lizenzen hat
- Die Herausforderung war dass die Daten aus verschiedenen Ländern aus unterschiedlichen <sup>Sources</sup> Quellen kamen und nicht im gleich format waren.

Derzeit arbeite ich bei Bosch als Praktikant im E-Bike Labor

Meine Hauptaufgabe ist es, NVH Daten von Ebike Sensoren zu sammeln und aufzubereiten

Ich ~~sammle~~ sammle und bereinige die Daten mit Python und Pandas. Anschließend präsentiere ich die Ergebnisse in Power BI - Dashboards

Um das zu tun, verarbeitet und zu bereinigen.

Danach füge ich die bereinigten Daten in eine eigene Datenbank

Aus der Datenbank hole ich die Daten und nutze sie in Power BI für Berichte und Dashboards.

Diese Bilder zeigen mir Ihnen die Ergebnisse

# Data Analyst (Bosch)

Ich arbeite im DataScience Team in der Ebike abteilung bei Bosch

Ich sammle NVH-Daten von Sensoren die bei Test der E-Bike entstehen

Ich habe eine Datenpipeline in Python entwickelt, um daten zu sammeln, zu verarbeiten und zu bereinigen

process clean

Danach lade ich die bereinigten Daten in eine MySQL Datenbank

Aus der datenbank hole ich die Daten und nutze sie in Power BI für Berichte und Dashboard

Diese Berichte werden im Power BI service <sup>published</sup> veröffentlicht, damit das Management sehen und Entscheidungen treffen kann.

# Warum möchten Sie bei uns arbeiten?

Ich möchte bei Ihnen arbeiten, weil ich sehe, dass Ihr Unternehmen datengetrieben arbeitet und moderne Technologien einsetzt

## Was sind Ihre Stärken?

Ich arbeite sehr struktuiert, lerne schnell und kann komplexe Daten sehr gut analysieren und verständlich darstellen

Außerdem kann ich sowohl selbstständig als auch im Team gut arbeiten

## Was sind Ihre Schwächen

Manchmal bin ich zu perfektionistisch und versuche etwas absolute richtige machen auch wenn es nicht immer notwendig ist  
Ich arbeite aber aktiv daran.

Außerdem sage ich manchmal zu schnell "Ja" zu zusätzlichen Aufgaben

## Wo sehen Sie sich in 3-5 Jahren

In den nächsten Jahren sehe ich mich als Data Analyst oder engineer mit tiefem Wissen in Datenmodellierung, Machine Learning und Cloud Technologien - Ich möchte mehr Verantwortung übernehmen, eigene Projekt und aktiv dazu beitragen

## Warum sollten wir Sie einstellen?

Sie sollten mich einstellen, weil ich praktische Erfahrung mit Daten habe. Ich kann Daten sammeln, bereinigen, analysieren und in Power BI darstellen.

Durch meine Arbeit bei Bosch, Siemens und meinen anderen Projekten kenne ich den ganzen Prozess sehr gut.

Ich lerne schnell, ich bin ~~immer~~ zuverlässig und finde immer eine gute Lösung für technische Aufgaben

## Arbeiten Sie lieber allein oder im Team?

Beides ich kann sehr gut im Team arbeiten. Außerdem ich kann allein arbeiten. Es kommt darauf an, was für die Aufgabe benötigt wird

## Erzählen Sie von einem Projekt, auf sie stolz sind.

In vielen meiner Power BI Projekte gibt es Performance Problem

Die Hauptgründ dafür sind

- Ein ineffizientes Datenmodell
- Falsche Spalten für Joins
- Zu viele unnötige Daten im Modell

Vertibag,

Comparisons using int (number type)

# Power BI Optimierung

- Nur benötigte necessary Spalten columns laden
- Nur relevante Zeilen rows laden
- Beziehungen zwischen Tabelen auf numerische Keys
- Bidirectional Beziehungen nur bei bedarf requirement verwenden
- Im datenmodel wurden alle Daten geladen, obwohl daten letzten zwei Jahre erforderlich required waren.

Ich habe nur noch meine Masterarbeit übrig  
das ist aber kein Problem. Ich kann das  
später machen oder ich kann auch das im  
abends oder am Wochenende arbeiten.

Ich kann auch für 6 Monat als Werkstudent  
arbeiten und in dieser Zeit meine Masterarbeit  
abschließen. Danach kann ich Vollzeit arbeiten

Ich lerne Deutsch und habe derzeit das  
Niveau B2 erreicht.

Momentan arbeite ich an meinen sprachfähigkeiten

Ich habe am Anfang eine Frage  
Sind Deutsch Kenntnisse für diese Position  
notwendig?



Letzte Woche habe ich meine  
Microsoft Fabric Data Engineering Zertifikat  
Prüfung bestanden

- Bei meiner aktuellen Arbeit bei Bosch habe ich einer Kollegin bei einem Machine learning Projekt geholfen
  - Ich habe die Daten mit Python (Pandas) gereinigt und vorbereitet damit sie diese für ihre Modelle benutzen konnte
  - Außerdem habe ich sie dabei ~~██████████~~ unterstützt die Ergebnisse der Modelle gut verständlich in Power BI darzustellen
- 
- Ich habe keine berufliche Erfahrung mit Azure DevOps aber ich benutze Fabric Deployment pipelines
  - Bei Bosch arbeiten wir mit Jira für Tickets und GitHub für Versionsverwaltung

Ich möchte bei Ihnen arbeiten, weil ich sehe, dass Ihr Unternehmen datengetrieben<sup>data driven</sup> arbeitet und moderne Technologien einsetzt<sup>use</sup>

Ist Deutsch für diese Rolle unbedingt notwendig<sup>absolute necessary</sup> oder ist es auch möglich auf Englisch zu sprechen?

Ich bin derzeit<sup>currently</sup> auf dem Deutsch Niveau<sup>level</sup> B2

- Reinigung und Transformation großer Datensätze waren eine ersten Herausforderung
- Daten waren oft unvollständig oder falsch formatiert
- Lösung: Nutzung von Power Query um Daten zubereinigen, Spalten zu trennen/Kombinieren und fehlende Werte zu behandeln
- Erstellung eines guten Datenmodells war schwierig
- Beziehungen zwischen Tabellen waren nicht immer klar
- Analyse der Datenquellen und Aufbau eines Sternschemas
- DAX Formeln waren am Anfang Komplex
- Übung, Lesen der Dokumentation und Testen verschiedner Ansätze
- Langsam Power BI Reports waren ein Problem
- Entfernen unnötiger Spalten, Nutzung von Aggregationen und DAX Optimierung.

- Das ist unser Arbeitsablauf
- Es besteht aus 5 Schritten
- NVH Ingenieure testen verschiedene "drive units" und zeichnen Messwerte mit Sensoren auf
- Datenqualität wird mit einem Python Script überprüft.  
und nach der Qualitätsprüfung werden die korrekten Daten in einem separaten Ordner gespeichert
- Schritt 3 ist ETL  
Die Daten aus diesem Ordner werden per ETL in die NVH-Datenquelle geladen  
ETL besteht aus drei Schritte
  - 1) Data Aquisition
  - 2) Data Processing
  - 3) Data Loading
- In Schritt 4 werden die Daten in der Datenbank geordnet  
und in Schritt 5 ich benutze diese Daten für Power BI Reports zu erstellen

## ETL

Zurzeit diese ETL - Prozess hat einige Problem

- ① Python code ist nicht gut
- ② Der Logik zur Daten Überprüfung funktioniert nicht zu 100 Prozent
- ③ Der Prozess ist nicht automatisiert

Topic: Masterarbeit

Wenn du dieses Projekt übernehmen möchtest deine Aufgaben wären

- 1) Python Code zu verbessern
- 2) Die ETL Pipeline irgendwie zu automatisieren

for Automatisieren kann ich ein Vorschlag machen

Du Kannst entweder Airflow verwenden oder Python notes in fabric einzusetzen