



Bachelorarbeit

im Studiengang

Wirtschaftsinformatik (EBIS)

zur Erlangung des Grades eines

Bachelor of Science

Thema

Umsetzung einer Business Intelligence (BI)-Lösung
als Demonstrator für die Lehre auf Basis von Open Source-
Daten in Google Cloud Plattform und Microsoft Azure

Vorgelegt von:

Yasein Abu Shikhada

geb. am 01.06.1997 in Taschkent

Maastrichter Ring 26, 60486, Frankfurt am Main

Matrikel-Nr.: 1303272 im Winter-Semester 2023/2024

am Fachbereich 3: Wirtschaft und Recht

der Frankfurt University of Applied Sciences

Erstprüfer: Dr. Jürgen Dietrich

Zweitprüfer: Professor Dennis Forster

Abgabedatum: 20.11.2023

Sperrvermerk

Die vorliegende Bachelorarbeit enthält vertrauliche Daten und ist daher besonders zu schützen.

Die Einsicht in der Bachelorarbeit ist Unbefugten nicht gestattet. Ausgenommen hiervon sind die Gutachter, sowie berechtigte Mitglieder des Prüfungsausschusses der Frankfurt University of Applied Sciences. Die Vervielfältigung und Veröffentlichung des Abschlussberichtes – auch auszugsweise – ist grundsätzlich nicht erlaubt.

Danksagung

„Ich möchte mich herzlich bei allen Professoren der Frankfurt University of Applied Sciences bedanken, bei denen ich die Ehre hatte, Vorlesungen zu besuchen und aus deren umfangreichem Wissen zu schöpfen. Die Zeit an der Universität war für mich besonders spannend, und die Lehrveranstaltungen haben mein Fachwissen nachhaltig geprägt.

Ein besonderer Dank gilt auch meinen Eltern, die mich auf meinem Studienweg stets unterstützt haben. Ihre Unterstützung war für mich von unschätzbarem Wert.

Besonders hervorheben möchte ich die Unterstützung von Professor Marcus Zinn. Seine Vorlesungen und die Gelegenheit, mit ihm nach den Vorlesungen zu sprechen, haben mir maßgeblich geholfen, meine Richtung im Bereich Wirtschaftsinformatik zu finden. Ohne seine Hilfe und Inspiration, die er mir geschenkt hat, hätte ich diese Arbeit möglicherweise nicht geschrieben.

Ich möchte mich herzlich bei Dr. Jürgen Dietrich bedanken für seine wertvolle Unterstützung in Wort und Tat.

Schließlich möchte ich mich auch bei mir selbst bedanken, für meine Disziplin und Neugier, die mir geholfen haben, dieses Studium erfolgreich zu absolvieren.

Vielen Dank an alle, die zu meinem akademischen Erfolg beigetragen haben.

Mit freundlichen Grüßen, Yasein Abu Shikhada“

Inhaltsverzeichnis

- I. Abbildungsverzeichnis
- II. Abkürzungsverzeichnis
- III. Inhalt
- IV. Eidesstattliche Erklärung
- V. Quellenverzeichnis

1. Kapitel 1: Einführung

1.1. Hypothese.....	9
1.2. Forschungsfrage.....	9
1.3. Einleitung.....	9
1.4. Warum ist es wichtig, das Verbraucherverhalten zu verstehen?.....	10
1.5. Business Intelligence.....	11
1.6. Business Intelligence in Planung, Reporting & Kontrolle (BI-supported controlling).....	12
1.7. BI Strategie und Unternehmensausrichtung.....	14
1.8. DWH, Data mining, OLAP und ETL.....	14
1.9. Business Intelligence Architekturen.....	16
1.9.1. Stove-Pipe-Ansatz.....	16
1.9.2. Core Data Warehouse Ansatz.....	17
1.10. Business Intelligence Plattformen.....	18

2. Kapitel 2 : Entwicklung einer BI Strategie

2.1. Technologische Ziele.....	20
2.2. Organisatorische Ziele.....	21
2.3. Ganzheitliches Vorgehensmodell zur Entwicklung einer BI-Strategie.....	22

3. Kapitel 3 : Data Governance

3.1. Data Governance und Datenethik.....	26
3.2. Privatsphäre ist wichtig.....	27

4. Kapitel 4 : Anforderungsanalyse & Praktische Umsetzung

4.1. Anforderungsanalyse & Projektplanung.....	29
4.1.1. Projektziel.....	29
4.1.2. Aufgaben von der Abteilung für Kundenentwicklung:.....	29
4.1.3. Ergebnisse und Metriken.....	29
4.1.4. Erfolgskriterien.....	29
4.2. Praktische Umsetzung.....	30
4.2.1. Datenquellen.....	30
4.2.2. Datensammlung, Datenmodellierung und ETL-Prozesse.....	31
4.2.2.1. Arten von Schemata (Stern und Schneeflocke).....	31
4.2.3. Google Cloud Plattform + Tableau.....	34
4.2.3.1. Google DataFlow.....	34
4.2.3.2. Google Big Query.....	35
4.2.3.3. SQL-Abfrage zur Erstellung von Zieltabelle (für den ganzen Zeitraum).....	36
4.2.3.4. SQL-Abfrage zur Erstellung von Zieltabelle (für Sommer Saison 2022).....	39
4.2.3.5. Datenvisualisierung & Gewinnung von Erkenntnissen.....	42
4.2.3.5.1. Datenvisualisierung, /-Analyse für den Zeitraum (2018-2022).....	42

4.2.3.5.2. Datenvisualisierung, /-Analyse für die Sommersaison 2022.....	51
4.2.4. Microsoft Azure + Power BI.....	53
4.2.4.1. Azure Analysis Services.....	53
4.2.4.2. Datenvisualisierung & Gewinnung von Erkenntnissen.....	62
4.2.4.2.1. Datenvisualisierung, /-Analyse für den Zeitraum (2018-2022).....	62
4.2.4.2.2. Datenvisualisierung, /-Analyse für die Sommersaison 2022.....	71
 5. Kapitel 5 : Auswertung der Visualisierungstechnologien und Ergebnis der Forschung	
5.1. Auswertung der Ergebnisse aus den Visualisierungen & Antworten auf Geschäftsfragen.....	74
5.2. Auswertung beider Visualisierungstechnologien : Power BI und Tableau.....	78
5.2.1. Microsoft Power BI.....	78
5.2.1.1. Stärken.....	78
5.2.1.2. Schwächen.....	80
5.2.2. Tableau.....	81
5.2.2.1. Stärken.....	81
5.2.2.2. Schwächen.....	82
5.3. Wesentliche Unterschiede „Tableau“ vs „Power BI“ aus Benutzersicht.....	82
5.4. Ergebnis der Forschung.....	83
5.5. Handlungsempfehlungen.....	84
5.6. Fazit.....	85

I. Abbildungsverzeichnis

Abbildung 1 : Einordnung der Verständnisweisen zu BI.....	15
Abbildung 2 : Stove-Pipe-Ansatz.....	16
Abbildung 3 : Core Data Warehouse Ansatz.....	18
Abbildung 4 : Magisches Quadrat der Business Intelligence Plattformen.....	19
Abbildung 5 : Perspektiven einer BI-Strategie.....	23
Abbildung 6 : Ganzheitliches Vorgehensmodell zur Entwicklung der BI Strategie.....	24
Abbildung 7 : Schneeflockenschema-Datenmodell für erste Zieltabelle.....	32
Abbildung 8 : Schneeflockenschema-Datenmodell für zweite Zieltabelle.....	33
Abbildung 9 : Google Cloud Plattform Architektur – ETL.....	35
Abbildung 10 : SQL-Abfrageprozess in BigQuery – Google Cloud Plattform.....	38
Abbildung 11 : SQL-Abfrageprozess in BigQuery – Google Cloud Plattform.....	41
Abbildung 12 : Jahresvergleich: Fahrradnutzung und durchschnittliche Tripdauer.....	42
Abbildung 13 : Verteilung der Nutzertypen (Subscriber vs. Customer).....	43
Abbildung 14 : Verteilung der Anzahl der Fahrten basierend auf das Geschlecht und Alter.....	44
Abbildung 15 : Häufigkeit und Nutzertypen der Startstationen nach Stadtgebiet und Stadtteil.....	45
Abbildung 16 : Häufigkeit und Nutzertypen der Endstationen nach Stadtgebiet und Stadtteil.....	47
Abbildung 17 : Einfluss der Niederschlagswahrscheinlichkeit auf die Anzahl der Fahrten.....	48
Abbildung 18 : : Korrelation zwischen durchschnittlicher Tagestemperatur und Anzahl der Fahrten.....	49
Abbildung 19 : Beziehung zwischen durchschnittlicher Windgeschwindigkeit (m/s) und Anzahl der Fahrten.....	50
Abbildung 20 : Darstellung der Startstationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken.....	51
Abbildung 21 : Darstellung der Abgabestationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken.....	52
Abbildung 22 : Einrichtung des „Cyclistic“-Verzeichnisses in Microsoft Analysis Services.....	54

Abbildung 23 : Einrichtung des „Cyclistic“-Verzeichnisses in Microsoft Analysis Services.....	56
Abbildung 24 : Einrichtung des „Cyclistic“-Verzeichnisses in Microsoft Analysis Services.....	58
Abbildung 25 : Integration von Daten aus Azure Blob Storage in Power BI.....	60
Abbildung 26 : Jahresvergleich: Fahrradnutzung und Tripdauer.....	62
Abbildung 27 : Verteilung der Nutzertypen (Subscriber vs. Customer).....	63
Abbildung 28 : Verteilung der Anzahl der Fahrten basierend auf das Geschlecht und Alter.....	64
Abbildung 29 : Häufigkeit und Nutzertypen der Startstationen nach Stadtgebiet und Stadtteil.....	65
Abbildung 30 : Häufigkeit und Nutzertypen der Endstationen nach Stadtgebiet und Stadtteil.....	67
Abbildung 31 : Einfluss der Niederschlagswahrscheinlichkeit auf die Anzahl der Fahrten.....	68
Abbildung 32 : Korrelation zwischen durchschnittlicher Tagestemperatur und Anzahl der Fahrten.....	69
Abbildung 33 : Beziehung zwischen durchschnittlicher Windgeschwindigkeit (m/s) und Anzahl der Fahrten	70
Abbildung 34 : Darstellung der Startstationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken.....	71
Abbildung 35 : Darstellung der Abgabestationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken.....	72

II. Abkürzungsverzeichnis

BI : Business Intelligence

BPM : Business Process Management

CSV : Comma-separated values

DAX : Data Analysis Expressions

DBMS : Datenbankmanagementsystem

DSGVO : Datenschutzgrundverordnung

DSS : Decision Support Systems

DWH : Data Warehouse

DWII : Data Warehouse in a Box

EIS : Executive Information Systems

ETL : Extrahierung – Transformation – Ladung

FIS : Führungsinformationssysteme

GCP : Google Cloud Plattform

GDPR : General Data Protection Regulation

MIS : Management Information Systems

NOAA : National Oceanic and Atmospheric Administration

NYC : New York city

OLAP : Online Analytical Processing

OLTP : Online Transactional Processing

ORM : Original Equipment Manufacturer

S/CAL : Server/Client Access License

SaaS : Software-as-a-Service

SQL : Structured query language

SSOT : Single Source of Truth

VizQL : Visual query language for databases

Kapitel 1 : Einführung

Im ersten Kapitel dieser Bachelorarbeit werden grundlegende Begriffe im Zusammenhang mit Business Intelligence eingeführt und erläutert. Die Forschungsfrage, die den Rahmen für die gesamte Arbeit bildet, wird ebenfalls prägnant dargestellt, um den Leser auf das Kernthema einzustimmen.

1.1. Hypothese :

Implementierung eines Business-Intelligence-Ansatzes wird dazu führen, dass Führungskräfte bessere und datengestützte Entscheidungen treffen, wenn sie auf eine umfassende Datenanalyse eines realen Unternehmensdatensatzes zurückgreifen.

1.2. Forschungsfrage:

Unter Verwendung von Datenanalyse und BI-Technologien wird in dieser Arbeit folgende Forschungsfrage untersucht:

Inwieweit führt die Anwendung des Business Intelligence Ansatzes mit Microsoft Azure und Power BI zur Verbesserung des Verständnisses der Fahrradnutzung und im Vergleich zur Verwendung von Google-Technologie in Kombination mit Tableau für die Analyse von Cyclistic-Daten in der Stadt New York?

1.3. Einleitung

Große Unternehmen erleben häufig Misserfolge bei der Einführung neuer Produkte, wie z. B. bei New Coke, Toshiba HD-DVD, VW Phaeton oder Microsoft Zune. Es ist durchaus üblich, dass etwa 60 bis 80 Prozent der neu eingeführten Produkte ein Jahr nach ihrer Markteinführung nicht mehr erhältlich sind.¹ Also, warum scheitern Produkte meistens bei der Einführung auf den Markt?

Das Fehlen einer Verbraucherorientierung in den Unternehmen ist eins der Faktoren, der zu diesem Problem beiträgt. Umfangreiche Forschungsarbeiten, darunter Studien von Srinivasan & Hanssens (2009) und Steiner (2009), haben durchweg die positiven Auswirkungen der Verbraucherorientierung auf Umsatz und Gewinn nachgewiesen. Folglich ist es für die Unternehmensleitung unerlässlich, die Ausrichtung ihres Geschäftsmodells auf die Kundenbedürfnisse zu einem vorrangigen Ziel zu machen.

Datenkompetenz, die Fähigkeit, Daten zu erstellen, zu verstehen und zu nutzen, stellt eine große Herausforderung für jedes Unternehmen und jede Organisation in der heutigen Zeit dar. Sie spielt eine entscheidende Rolle, wenn es darum geht, Menschen mit dem Wissen und den Fähigkeiten auszustatten, die für die Entscheidungsfindung und kritisches Denken in einer Gesellschaft notwendig sind, in der die Wahrheit oft verzerrt wird. Das Verständnis und die Anwendung von Daten sind auch für kritische Situationen in der Geschäftswelt, wie: Krisenmanagement, Lieferkettenmanagement, Kundenbeziehungen, Sicherheitsvorfälle, Markttrends und Personalwesen von grundlegender Bedeutung.

¹ (vgl. Müller & Schroiff, 2013, S. 7)

Obwohl die Bedeutung der datengesteuerten Entscheidungen weithin anerkannt ist, hat man sich nicht in der Vergangenheit ausreichend auf die grundlegenden Techniken und Ansätze für eine effektive Verwaltung von Daten konzentriert, unabhängig davon, ob es sich um Big Data oder Little Data handelt. Die Verwendung visueller Skriptsprachen und Software-Plug-ins hat die Aufmerksamkeit, die der Datenanalyse gewidmet wird, überschattet, nicht nur im Bereich der Architektur, sondern auch in anderen Branchen. Die Verbesserung der Datenkompetenz und das Verständnis für die Rolle der Datenanalyse als entscheidender Aspekt einer kritischen Praxis sind Herausforderungen, denen sich die Geschäftswelt stellen müssen. Die Wirtschaft hat anschließend das Potenzial von Daten für die Entscheidungsfindung erkannt und benutzerfreundliche Tools entwickelt, die es nicht nur Architekten, sondern auch anderen Menschen, die über begrenztes Fachwissen im IT-Bereich verfügen, ermöglichen, Daten abzufragen und komplexe Analysen durchzuführen, ohne dass umfangreiche Kenntnisse in Skriptsprachen oder Codierung erforderlich sind.

Der Begriff "Business Intelligence" wurde erstmals 1958 in dem vom Informatiker H.P. Luhn veröffentlichten Aufsatz "A Business Intelligence System" eingeführt.² Schon in den 1950er Jahren sah Luhn die Bedeutung der Echtzeit-Datenanalyse und des Data Mining für die Unterstützung von Geschäftsentscheidungen voraus. Als die Unternehmenswelt in den folgenden Jahrzehnten stark in die Datenanalyse investierte, gewöhnten sich die Führungskräfte daran, relevante Informationen zur Hand zu haben, die klar, zuverlässig und manipulierbar waren, um Antworten auf unterschiedliche relevante Fragen zu finden, die sich aus der Datenanalyse ergaben.

1.4. Warum ist es wichtig, das Verbraucherverhalten zu verstehen?

Kundenorientierung ist entscheidend für den Erfolg eines Unternehmens, insbesondere auf den heutigen Käufermärkten. In der Vergangenheit hatten die Verkäufer auf einem Verkäufermarkt, auf dem die Nachfrage das Angebot überstieg, die Oberhand. Sie konnten ihre Produkte verkaufen, ohne sich auf bestimmte Kundengruppen zu konzentrieren. Aufgrund des technologischen Fortschrittes hat sich die Marktdynamik im Laufe der Jahre verändert. Die Marktsättigung in verschiedenen Branchen und des Eindringens ausländischer Wettbewerber durch die Globalisierung beeinflussten wesentlich den Markt. Diese Faktoren haben die Märkte in käuferorientierte Märkte verwandelt, in denen die Verbraucher aus einer breiten Palette ähnlicher Angebote wählen können. Daher mussten die Unternehmen ihren Ansatz von einer Vertriebs- und Produktionsausrichtung auf eine Ausrichtung auf den Kunden umstellen. Diese veränderte Ausrichtung spiegelt die veränderte Marktdynamik und die Bedeutung des Verständnisses und der Erfüllung von Kundenbedürfnissen wider.³ Somit stieg die Konkurrenz an, was im Endeffekt zu einem starken Wettbewerb auf dem Markt geführt hat.

² (vgl. Chamoni/Gluchowski 2010, S. 18)

³ (Gelbrich et al., 2018; Meffert et al., 2019, S. 13)

Um auf Käufermärkten erfolgreich zu agieren, bedarf es einer Abgrenzung von Konkurrenten und der Ausstattung von Produkten mit Eigenschaften, die eine spezifische Zielgruppe ansprechen. Hierfür ist es notwendig, die Zielgruppe und den Verbraucher sehr gut zu kennen und seine Bedürfnisse zu verstehen, um ihn die Lösung in Form von einem Produkt anbieten zu können. Die Vorteile, die das Wissen über das Verbraucherverhalten für das Marketing mit sich bringt, lassen sich anhand der sogenannten "4Ps" verdeutlichen ⁴: Die Kenntnis des Konsumentenverhaltens ermöglicht unter anderem eine an die Bedürfnisse der Zielgruppe angepasste Produktentwicklung und -gestaltung (Produktpolitik), die Ausnutzung der maximalen Zahlungsbereitschaft des Konsumenten (Preispolitik), die an die Präferenzen des Konsumenten angepasste Distribution des Produktes (Place-Politik) sowie die gezielte Ansprache der Konsumentenbedürfnisse über verschiedene Kommunikationskanäle wie TV-Werbung, Online-Werbung, Sponsoring etc. (Promotionspolitik).⁵

1.5. Business Intelligence

Experten und Forscher haben dem Begriff Business Intelligence verschiedene Definitionen zugeordnet. Diese Disziplin besteht aus einer Palette von Aktivitäten und Prozessen, vor allem die Datenerfassung, tiefgehende Analysen, die Erstellung von fundierten Berichten und die Gestaltung von aussagestarken Dashboards und Visualisierungen. Das ultimative Ziel von BI ist es, Unternehmen mit wertvollen und stets aktuellen Informationen zu versorgen, um ihnen dabei zu helfen, kluge Entscheidungen zu treffen und ihre Geschäftsergebnisse auf eine neue Ebene zu heben. ⁶

BI ist der geheime Verbündete jeder modernen Unternehmensführung. Es ist eine beeindruckende Sammlung von Technologien, Werkzeugen und Methoden, die wie ein kraftvolles Werkzeugset darauf abzielen, Unternehmen bei ihrer Entscheidungsfindung und strategischen Planung zu unterstützen. BI Tools sind in der Lage, Informationen aus einer Vielzahl von Quellen zu sammeln und nahtlos zu integrieren, um wertvolle Einblicke in die Performance eines Unternehmens zu gewinnen.⁷ Hierbei werden Trends und Muster aufgespürt, Prognosen erstellt und Risiken minimiert. Und das Ganze erfolgt in einer interaktiven Form.

Es ist klar, dass Data Analyse und Business Intelligence viele Gemeinsamkeiten aufweisen. Die Strukturierung des Themas Data Analytics im Unternehmen orientiert sich an der dargestellten Unterteilung in "Deskriptive Analytics" (die Analyse von Daten aus der Vergangenheit), "Prädiktive Analytics" (die Vorhersage von zukünftigen Ereignissen) und "Präskriptive Analytics" (die Ableitung von Handlungsempfehlungen für das Unternehmen). Oft wird "Deskriptive Analytics" auch mit den im Unternehmen etablierten Begriffen "Business Intelligence und Reporting" gleichgesetzt. Zur besseren Differenzierung werden "Prädiktive" und "Präskriptive Analytics" zusammen als "Advanced Analytics" bezeichnet. Organisatorisch sind "Business Intelligence" und "Advanced Analytics" in einer zentralen Abteilung gebündelt, die der Group IT zugeordnet ist.⁸

⁴ (vgl. Trommsdorff & Teichert, 2011, S. 14)

⁵ (vgl. Hoffmann/Akbar 2023, S. 6)

⁶ (vgl. Azvine et al. 2005, S. 15)

⁷ (vgl. Negash 2004, S. 18)

⁸ (vgl. Seiter 2023, S. 152)

1.6. Business Intelligence in Planung, Reportwesen & Kontrolle (BI-supported controlling)

Basierend auf den Unternehmenszielen waren die Planung und Berichterstattung schon immer entscheidende Instrumente für das Management und Controlling, um Koordination und Kommunikation sicherzustellen. Es ist bemerkenswert, dass in den letzten Jahren vermehrt Beiträge und Bücher veröffentlicht wurden, die sich entweder intensiver mit der Planung oder der Berichterstattung beschäftigen. Darüber hinaus existieren Publikationen über praktische Lösungen, die speziell die Besonderheiten, Probleme und Umsetzungsmöglichkeiten sowohl von Berichterstattungs- als auch von Planungssystemen beleuchten. Aus diesem Grund wird diese Studie systematisch den Aspekt der Integration beider Seiten untersuchen, beginnend mit der Definition, den Zielen und Aufgaben der Planung sowie der Berichterstattung.

In vielen Fällen gelten die Aussagen sowohl für die Planung als auch für die Berichterstattung gleichermaßen. Falls es Unterschiede zwischen Planung und Berichterstattung gibt, werden diese besonders hervorgehoben.⁹

Reporting bezeichnet, wie man sich vorstellen kann, verschiedene papier- oder IT-basierte Managementberichte und Analysen, in denen unterschiedliche Arten von Informationen mithilfe von Tabellen, Diagrammen und Kommentaren präsentiert werden.¹⁰ Es gibt viele verschiedene Definitionen von Reporting. Zum Beispiel versteht man im Gabler Wirtschaftslexikon Reporting als alle systematisch erstellten Informationen, die in schriftlicher oder elektronischer Form für interne und externe Adressaten relevant sind.¹¹ Horvath hingegen definiert das Management-Reporting viel prägnanter. Als Teil des operativen Managementsystems umfasst das Management-Reporting nach Horvath die Phasen der Informationsbereitstellung und -übermittlung sowie der Informationsnutzung.¹² Mit der Einführung eines Data Warehouses im BI-unterstützten Controlling wird eine Informationsstufe geschaffen, die zur Verbesserung der Entscheidungsqualität des Managements im Sinne eines Single Point of Truth beiträgt.¹³ Die Bezeichnung für die einzige Wahrheit wird in der Regel als "Single Point of Truth" oder "Single Source of Truth" (SSOT) verwendet. Dieser Begriff beschreibt das Konzept, dass es in einem bestimmten Kontext oder einer bestimmten Organisation eine zentrale, autoritative und unbestreitbare Quelle für bestimmte Daten oder Informationen gibt, auf die alle Benutzer vertrauen können.

Definition und Bedeutung von Reporting kann je nach Kontext variieren, aber es ist unbestreitbar ein entscheidender Aspekt des Informationsmanagements in Unternehmen, da es dazu beiträgt, fundierte Entscheidungen zu treffen und die Transparenz in Bezug auf Geschäftsdaten zu gewähren.

⁹ (vgl. Schön 2023, S. 39)

¹⁰ (vgl. ebd.: S. 17)

¹¹ (Gablerlexikon 2004, S. 364)

¹² (vgl. Gleich 2008, S. 18-20)

¹³ (vgl. Schön 2023, S. 17)

Zur Beurteilung von Leistung, Produktivität und Rentabilität sind Orientierungs- und Vergleichsmaßstäbe in allen Unternehmensbereichen erforderlich. Eine zukunftsorientierte Unternehmensführung benötigt entscheidungsrelevante Informationen mit prognostischem Charakter. Dies gilt für alle Teilbereiche der Unternehmensplanung. Gemäß Wilds Definition wird Unternehmensplanung wie folgt definiert:

Unternehmensplanung umfasst das systematische, zukunftsorientierte Denken und Festlegen von Unternehmenszielen sowie Maßnahmen, Mitteln und Wegen zu ihrer Erreichung.¹⁴

Die Unternehmensplanung kann nach den Kriterien des Planungsobjekts, des Planungsinhalts und des Planungshorizonts differenziert werden. Objekte sind die betrachteten Subjekte und Verantwortungsbereiche, die dem Planungsprozess zugrunde liegen, wie Unternehmen, Kostenstellen, Kostenträger, Kunden, Lieferanten, Mitarbeiter, Einrichtungen usw. Der Inhalt umfasst die Mengen- und Wertstrukturen sowie andere Eigenschaften, die in Bezug auf die Planungsobjekte prognostiziert und budgetiert werden.¹⁵

Während Business Intelligence eher durch die Verarbeitung von Daten und Informationen gekennzeichnet ist, fasst das Business Performance Management die Interaktion zwischen dem Informationssystem und den primären und sekundären Managementfunktionen zusammen. BPM integriert daher Informationstechnologie und Geschäftsmethoden zur Steuerung und Kontrolle eines Unternehmens¹⁶. Durch Business Intelligence unterstütztes Controlling ist die zielgerichtete Entscheidungsunterstützung für die Führungskräfte im Unternehmen und hilft bei Aufgaben wie Planung, Analyse, Kontrolle und Koordination. Dabei werden Informationstechnologien wie Data Warehousing und Big Data-Technologien genutzt.¹⁷

„Die durch Business Intelligence unterstützten Controlling-Services dienen dazu, die Entscheidungsgrundlagen für das Management zu verbessern, die erforderliche Transparenz für unternehmerisches Handeln zu schaffen und Beziehungen zwischen zuvor isolierten Informationen aufzuzeigen“¹⁸. In dieser Hinsicht stellt das durch Business Intelligence unterstützte Controlling eine Aufklärung des unterstützenden Charakters der Business Intelligence für das Controlling dar, wobei der Einsatz neuer Informationstechnologien mit den professionellen Anwendungsmethoden des Controllings für das Management verknüpft wird.¹⁹

Da eine gleichzeitige Gesamtgeschäftsplanung, die die Berücksichtigung zahlreicher komplexer Geschäftsbeziehungen erfordert, als nicht realisierbar angesehen wird, ergibt sich eine kohärente und konsistente Unternehmensplanung lediglich aus aufeinanderfolgenden Business-Support-Planungen.²⁰

¹⁴ (vgl. Wild 1981, S. 13)

¹⁵ (vgl. Schön 2023, S. 39)

¹⁶ (vgl. ebd.: S. 454)

¹⁷ (vgl. ebd.: S. 455)

¹⁸ (vgl. Schrödl 2009 S. 13)

¹⁹ (vgl. Schön 2023, S. 456)

²⁰ (vgl. Scherrer 1991, S. 143)

1.7. BI Strategie und Unternehmensausrichtung

Die Unternehmensstrategie bildet die grundlegende Orientierung für die Entwicklung eines Unternehmens. Das Ziel der Strategie besteht darin, den langfristigen Erfolg des Unternehmens zu gewährleisten.²¹ In diesem Zusammenhang wird die Auffassung vertreten, dass es von großer Bedeutung ist, eine spezielle Business Intelligence (BI)-Strategie zu entwickeln, um die Lücke zwischen der Unternehmensstrategie und den Strategien auf Bereichs- oder Funktionsebene erfolgreich zu schließen. Indem direkte Impulse aus der Unternehmensstrategie einfließen, wird sichergestellt, dass BI grundsätzlich zur strategischen Ausrichtung des Unternehmens beiträgt. Dies kann beispielsweise die Ziele in Bezug auf Innovationen, Alleinstellungsmerkmale, die Art des (internationalen) Wachstums, die gewünschte Unternehmensstruktur oder die Steuerungsmechanismen des Unternehmens umfassen. Durch eine gestufte Zielhierarchie können die Auswirkungen auf Business Intelligence konkretisiert werden.²²

IT-Strategie soll sicherstellen, dass alle IT-bezogenen Aktivitäten und sowohl operative als auch strategische Entscheidungen im Einklang mit den Gesamtzielen des Unternehmens stehen.²³ Die BI-Strategie kann ihrerseits Teilstrategien wie Reporting- oder Data-Warehouse-Strategien enthalten. Diese sind jedoch keine eigenständigen Strategien, sondern Teilaspekte der BI-Strategie darstellen.²⁴

Ein enges Verhältnis zur BI-Strategie hat die Datenqualitätsstrategie, da hohe Datenqualität eine grundlegende Voraussetzung für die Akzeptanz von BI darstellt. Idealweise verfügt ein Unternehmen über eine etablierte, eigenständige Datenqualitätsstrategie und ein daraus resultierendes Datenqualitätsmanagement über alle Systemgrenzen hinweg. In der Praxis werden Datenqualitätsprobleme in den operativen Systemen oft erst bei der Einführung von BI-Systemen offensichtlich.²⁵ Daher wird die Entwicklung einer Datenqualitätsstrategie häufig als Folgeprojekt einer BI-Strategie oder als Voraussetzung für neue BI-Initiativen betrachtet.

Die enge Verknüpfung zwischen Unternehmensstrategie, BI-Strategie und Datenqualitätsstrategie verdeutlicht die Notwendigkeit einer ganzheitlichen Herangehensweise bei der Implementierung von Business Intelligence. Dies trägt dazu bei, dass BI die strategischen Ziele des Unternehmens unterstützt und gleichzeitig die Qualität der verarbeiteten Daten sicherstellt.

1.8. DWH, Data Mining, OLAP und ETL

Management Support Systems (SS, deutsch Führungsunterstützungssysteme oder Führungsinformationssysteme) umfasst eine Vielzahl von analytischen Systemen. Management Information Systems (IS), Decision Support Systems (DSS) und Executive Information Systems (EIS) sind Begriffe, die verschiedene Arten analytischer Informationssysteme umfassen, deren Entwicklung bis in die 1970er Jahre zurückreicht.

²¹ (vgl. Hungenberg 2004, S.4)

²² (vgl. Gansor et al. 2010, S. 43)

²³ (vgl. Fröhlich/Galsner 2007, S.174)

²⁴ (vgl. Bauer 2009, S. 380)

²⁵ (vgl. Otto u.a. 2008, S. 215)

In MIS sind Systeme, die das mittlere Management mit Berichten unterstützen, um tägliche, strukturierte Entscheidungen zu unterstützen. Diese Berichte basieren auf Unternehmensdaten aus den operativen Systemen, können teilweise historisiert sein und haben normalerweise einen engen Subjektbezug.²⁶

Das Data Warehouse ist ein zentrales Archiv für Daten in jedem Unternehmen. In diesem Prozess werden Daten aus verschiedenen Quellen gesammelt und in ein einheitliches Gesamtbild integriert.²⁷

Das Data Mining ist die Verwendung statistischer Methoden und Algorithmen, um wertvolle Erkenntnisse aus den gesammelten Daten zu gewinnen.²⁸

Online Analytical Processing (OLAP) ist eine Software, die eine interaktive Datenanalyse in Rekordzeit ermöglicht, was eine dynamische Entscheidungsfindung unterstützt.

Die wichtigsten Darstellungskomponenten sind das Reporting und die Dashboards, die Analysen auf eine verständliche und visuell ansprechende Weise präsentieren.²⁹

Große Organisationen verfügen in der Regel über multidimensionalen Daten, d.h. Daten, die in mehreren Kategorien zur Nachverfolgung und Analyse unterteilt werden können. Somit ist OLAP laut IBM eine Software, die entwickelt wurde, um große Datenmengen aus einem Data-Warehouse, Datamart oder einem anderen zentralisierten Datenspeicher schnell und effizient multidimensional zu analysieren.³⁰

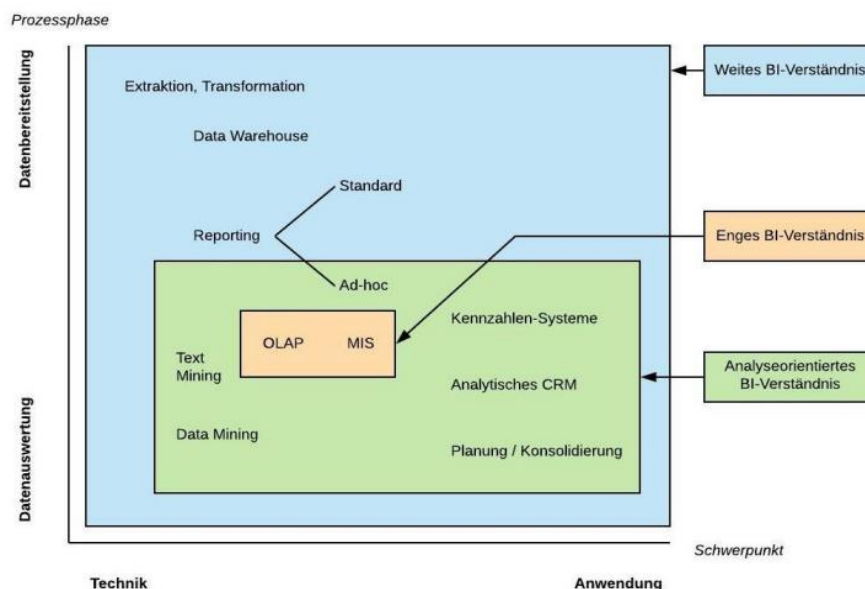


Abbildung 1: Einordnung der Verständnisweisen zu BI

(Quelle: Eigene Darstellung, in Anlehnung an Kemper, Baars und Mehanna (2010, S. 4))

²⁶ (vgl. Chamoni/Gluchowski 2010, S. 95)

²⁷ (vgl. Reinking 2023, S. 12)

²⁸ (vgl. ebd.: S.12)

²⁹ (vgl. Reinking 2023, S. 12)

³⁰ (vgl. Was ist OLAP? o. D.)

1.9. Business Intelligence Architekturen

1.9.1. Stove-Pipe-Ansatz

Der Stove-Pipe-Ansatz, auch unter dem Namen Ofenrohr-Ansatz bekannt, wird in Unternehmen in denjenigen Fällen angewandt, in denen keine übergreifenden Auswertungen benötigt sind.³¹

Dieser Vorgehensweise beinhalten die Extraktion und Verarbeitung von Informationen aus internen und externen Quellen, die speziell auf bestimmte Anwendungsgebiete oder Abteilungen zugeschnitten sind. Die Informationen werden in isolierte Datenmärkte übertragen, um den spezifischen Anforderungen einzelner Bereiche gerecht zu werden. In Situationen, in denen umfassende unternehmensweite Auswertungen nicht erforderlich sind, kann diese dezentrale Struktur mit unabhängigen Datenmärkten durchaus sinnvoll sein. Dennoch entstehen dabei isolierte Datenspeicher, da die Daten für jeden Bereich separat aus den Quellsystemen extrahiert und umgestaltet werden. Dies kann zu Duplikationen und potenziellen Unstimmigkeiten führen, insbesondere im Falle von erforderlichen Veränderungen.³²

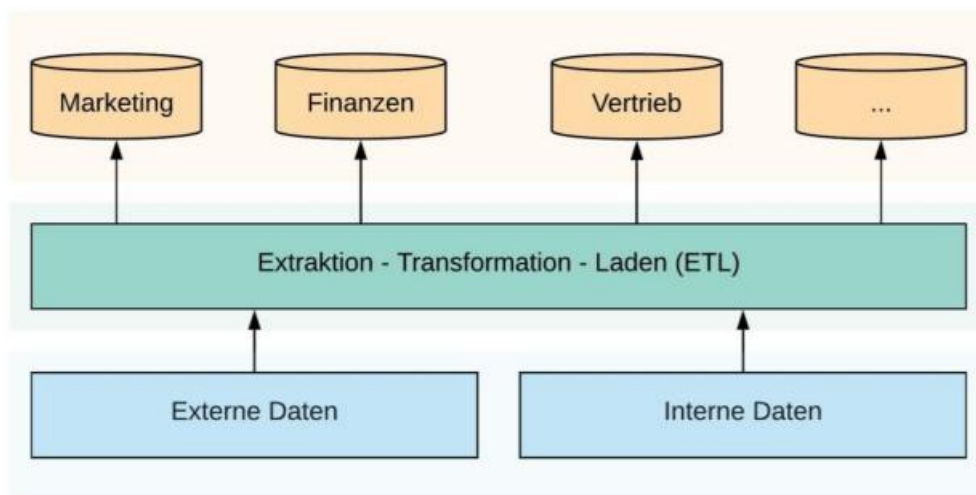


Abbildung 2 : Stove-Pipe-Ansatz

(Quelle: Business Intelligence Entwurf und prototypische Implementierung einer Omnichannel-Marketinganalyseplattform zur kanalübergreifenden Kampagnen-Steuerung auf Basis von Power BI ; 2020 ; Schneider)

³¹ (vgl. Hahne 2014, S. 26)

³² (vgl. ebd.: S. 26)

1.9.2. Core Data Warehouse Ansatz

Wenn analytische Anwendungen nur Daten aus einer bestimmten Anwendungsdomäne erfordern, kann der Verzicht auf Data Marts als eine mögliche Option in Betracht gezogen werden. Stattdessen sollte ein zentrales Core Data Warehouse eingerichtet werden, auf dem die Analysen direkt durchgeführt werden. Abgesehen von seiner Analysefunktion dient ein Core Data Warehouse auch der Datensammlung und Integration auf einem Speicherplatz. Dadurch wird die Qualität der Daten sichergestellt und eine Verteilungsfunktion gewährleistet.³³

Dieser Architekturansatz, wie in Abbildung 3 dargestellt, hat allerdings Einschränkungen hinsichtlich der Benutzeranzahl und der Handhabung größerer Datenvolumina, die direkt ausgewertet werden.³⁴ Aufgrund des Mangels an anwendungsspezifischer Verarbeitung dispositiver Daten ist dieser Ansatz weniger geeignet, wenn Daten aus verschiedenen Anwendungsdomänen integriert werden müssen, da eine kontextbezogene Aggregation von betriebswirtschaftlichen Daten fehlt.

Die zentrale Architektur eines Core Data Warehouse erleichtert den Betrieb und die Pflege des Systems im Vergleich zu einer Silo-Architektur. Aufgrund der geringen Anzahl von Komponenten sind Änderungen in den Datenstrukturen direkt für alle Anwendungen sichtbar. Dies ermöglicht eine schnelle Bereitstellung von Änderungen im Rahmen eines Change-Prozesses. Bei komplexeren Lösungen stößt dieser Ansatz jedoch aus Performance- und Administrationsgründen schnell an seine Grenzen.³⁵

In einem Core Data Warehouse werden dispositiv verarbeitete Daten nach den anfänglichen Transformationsschritten zur Bereinigung und Harmonisierung gespeichert und können für vielfältige Auswertungszwecke und Benutzer mit hoher Mehrfachverwendbarkeit und detaillierter Darstellung genutzt werden. Die Integrationsfunktion eines Core Data Warehouse ermöglicht Analysen auf harmonisierten Datenbeständen. Bei Geschäftsbereichen mit stark voneinander abweichenden Geschäftsprozessen stößt diese Integration jedoch schnell an ihre Grenzen, da eine umfassende Integration auf allen Ebenen und für alle Geschäftsbereiche oft nicht kosteneffizient durchführbar ist.³⁶

³³ (vgl. Hahne 2014, S. 28)

³⁴ (Sinz/Ulbrich-vom-Ende 2009, S. 188)

³⁵ (Kemper et al. 2010, S. 23)

³⁶ (vgl. Hahne 2014, S. 29)

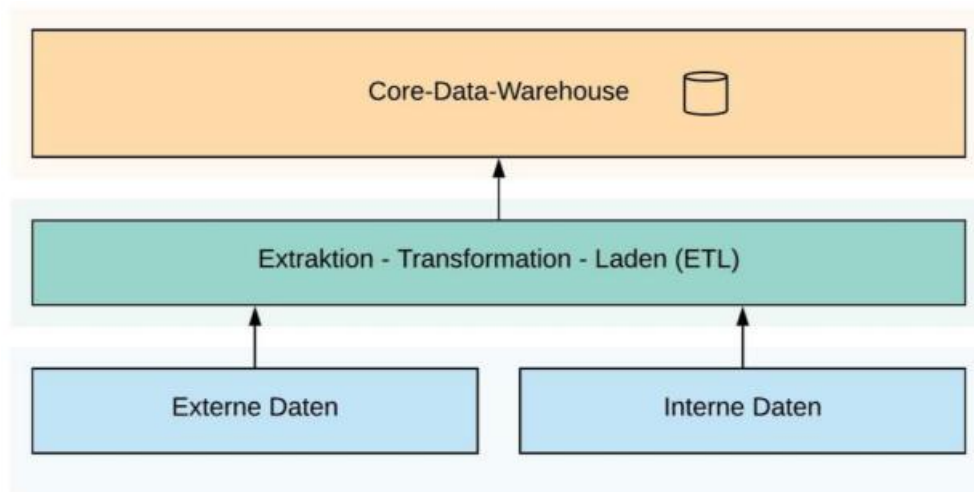


Abbildung 3: Core Data Warehouse Ansatz

(Quelle: Business Intelligence Entwurf und prototypische Implementierung einer Omnichannel-Marketinganalyseplattform zur kanalübergreifenden Kampagnen-Steuerung auf Basis von Power BI ; 2020 ; Schneider)

1.10. Business Intelligence Plattformen

Die Abbildung 4 zeigt, dass Tableau und Power BI als herausragende BI-Plattformen hervorstechen.

Laut Gartner sind Tableau und Power BI eindeutig die bevorzugten Tools für Business Intelligence. Sie befinden sich im Bereich „Market Leaders“. Diese sind angeblich die besten Plattformen für Unternehmen, die auf datengesteuerte Entscheidungen setzen und nach Exzellenz in der Datenanalyse streben.

Power BI ist ein Business Intelligence-Tool von Microsoft, einem der größten Technologieunternehmen der Welt. Tableau hingegen ist eine BI-Plattform, die von Tableau Software entwickelt wurde. Beide Plattformen bieten leistungsstarke Datenanalyse- und Visualisierungsfunktionen.

Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms



Abbildung 4 : Magisches Quadrat der Business Intelligence Plattformen
(Quelle : Richardson et al., 2020)

Kapitel 2 : Entwicklung einer BI Strategie

Im folgenden Kapitel werden sowohl die allgemeinen Ziele hinter einer BI-Anwendung beleuchtet als auch die Schritte zur Entwicklung einer nachhaltigen BI-Strategie, die zum Einsatz bereit sein kann.

2.1. Technologische Ziele

In der Vergangenheit wurde der technologische Fortschritt maßgeblich vorangetrieben, um die Anwendung von Business Intelligence zu verbessern. Beispielsweise wäre die Umsetzung des Data-Warehouse-Konzepts ohne die Verfügbarkeit großer Speicherkapazitäten nicht möglich gewesen. Ähnlich wird der Preisverfall von Hauptspeicherbausteinen die nächste Stufe für BI-Anwendungen markieren und die hochleistungsfähige Analyse großer Datenmengen ermöglichen.³⁷ Daher spielt der technologische Reifegrad eine entscheidende Rolle bei der effizienten Implementierung von BI-Systemen.

Im Streben nach technologischer Weiterentwicklung und Effizienz verfolgt Business Intelligence (BI) laut Tom Gansor et. al. mehrere technologische Ziele:

Optimierung der System-Performance, Integration neuester Technologien, Erweiterung von Hardware und Software, Datenarchivierung und -management und Serviceorientierte Implementierung von BI-Diensten.³⁸

Die Priorität dabei liegt auf der kontinuierlichen Optimierung der Leistungsfähigkeit der BI-Systeme, um sicherzustellen, dass schnelle Datenverarbeitung und effiziente Abfragefunktionen gewährleistet sind. Hierfür kommen moderne Technologien wie in-Memory-Analysesysteme und Data Warehouse-in-a-Box (DWII)-Appliances zum Einsatz. Diese Innovationen stärken die gesamte BI-Infrastruktur und bringen sie auf den neuesten Stand.

Sowohl die BI-Hardware als auch die -Software werden kontinuierlich erweitert und angepasst, um den Anforderungen von Echtzeitanwendungen gerecht zu werden. Dieser dynamische Ansatz sorgt dafür, dass die BI-Systeme mit den sich ständig ändernden Anforderungen und Herausforderungen Schritt halten können.

Darüber hinaus wird eine effektive Datenverwaltung sichergestellt. Ältere Daten werden gezielt in Archivierungssysteme wie Nearline Storage verschoben, um die Datenmenge in aktiven Systemen zu reduzieren. Dies führt nicht nur zu einer besseren Leistung, sondern erleichtert auch die effiziente Datennutzung und -pflege.

Die serviceorientierte Bereitstellung ausgewählter BI-Dienste ermöglicht den Nutzern eine flexible und skalierbare Nutzung. Diese Herangehensweise trägt zu einer effizienten Ressourcennutzung bei, was wiederum zu einer insgesamt gesteigerten Effizienz in der Datenverarbeitung und -analyse führt.

Diese technologischen Ziele sind die tragenden Säulen eines effizienten BI-Systems und spielen eine Schlüsselrolle dabei, Daten in wertvolle Erkenntnisse umzuwandeln, die als Grundlage für gut durchdachte Unternehmensentscheidungen dienen.

³⁷ (vgl. Gansor et al. 2010, S. 42)

³⁸ (vgl. ebd.: S. 42)

2.2. Organisatorische Ziele

Oftmals wird die Organisation von Business Intelligence erst nach der Einführung von BI-Systemen in den Fokus gerückt, was darauf zurückzuführen sein kann, dass in einigen Unternehmen BI, insbesondere die Implementierung eines Data Warehouses, zunächst lediglich als eine neue Art von IT-Systemen betrachtet wird, wobei traditionelle Entwicklungsmaßstäbe angewandt werden. Erst später wurde erkannt, dass nach der Implementierung von BI-Systemen ein anderer, den speziellen Anforderungen von BI gerecht werdender Maßstab angelegt werden muss.³⁹

Die Organisation und Berücksichtigung der speziellen Anforderungen von Business Intelligence sollte bereits in der Konzeptionsphase von BI-Systemen erfolgen, um einen reibungslosen Betrieb sicherzustellen.

Um die organisatorische Effektivität von Business Intelligence (BI) zu maximieren, werden verschiedene Ziele laut Tom Gansor et al. verfolgt:

Entwicklung einer BI-Governance, Verbesserung des Fachbereichs-IT-Zusammenspiels, Bündelung von Ressourcen und Know-how, Effizienzsteigerung im Betrieb von BI-Lösungen, Verkürzung der Entwicklungszyklen, Definition eines Standards, Dienstleistungsportfolios, Outsourcing definierter BI-Leistungen und Standardisierung von Betriebs- und Entwicklungsprozessen.⁴⁰

Die Etablierung einer effektiven BI-Governance-Struktur ist von zentraler Bedeutung, um sicherzustellen, dass Business Intelligence in einem Unternehmen effizient und zielführend eingesetzt wird. Diese Governance-Struktur schafft klare Verantwortlichkeiten und Prozesse, um sicherzustellen, dass BI-Lösungen den Geschäftsanforderungen entsprechen. Die enge Zusammenarbeit zwischen Fachbereichen und der IT ist ein Schlüsselfaktor, um sicherzustellen, dass die entwickelten BI-Lösungen tatsächlich die Bedürfnisse und Ziele des Unternehmens unterstützen.

Darüber hinaus bündelt die Zentralisierung von Ressourcen und Know-how in einer Einheit die Kräfte, um die Effizienz in der Entwicklung und Umsetzung von Business Intelligence zu steigern. Dies ermöglicht nicht nur eine schnellere Bereitstellung von BI-Lösungen, sondern fördert auch die konsistente Nutzung von bewährten Methoden und Technologien.

Die Implementierung optimierter Betriebsprozesse trägt dazu bei, die Effizienz bei der Bereitstellung von BI-Lösungen weiter zu steigern. Diese Prozesse beschleunigen die Entwicklung von BI-Anwendungen, was zu schnelleren Reaktionen auf geschäftliche Anforderungen führt.

Ein einheitliches Dienstleistungsportfolio für BI sorgt für klare Standards und Angebote, wodurch die gesamte Organisation einheitliche Maßstäbe und Qualitätskontrollen nutzen kann. In einigen Fällen kann die Auslagerung bestimmter BI-Aufgaben oder -Services dazu beitragen, Effizienz und Expertise zu nutzen, insbesondere wenn diese Aufgaben nicht zu den Kernkompetenzen des Unternehmens gehören.

³⁹ (vgl. Gansor et al. 2010, S. 41)

⁴⁰ (vgl. ebd.: S. 42)

Diese organisatorischen Ziele sind entscheidend, um sicherzustellen, dass Business Intelligence effektiv in die Unternehmensstruktur integriert ist und den Geschäftsanforderungen gerecht wird.

2.3. Ganzheitliches Vorgehensmodell zur Entwicklung einer BI-Strategie

Im Folgenden werden die grundlegenden Prinzipien zur Entwicklung einer Strategie im Bereich Business Intelligence (BI) diskutiert. In diesem Sinne orientiert sich die Strukturierung an den zentralen Aspekten einer BI-Strategie. Zuerst erfolgt die Betrachtung der Vision und der Ziele zur effizienten Steuerung des Unternehmens durch informationsbasierte Ansätze. Anschließend bildet das Konzept der Unternehmensführung das inhaltliche Fundament dieser BI-Strategie. Die konkrete Ausgestaltung der BI-Strategie erfolgt durch die Definition von Architektur, Anwendungen und Technologie ⁴¹. Dabei ist sowohl die technologische als auch die organisatorische Perspektive für das Unternehmen von höchster Bedeutung, insbesondere im Kontext von Business Intelligence.

Die Entwicklung einer Strategie ist grundsätzlich ein kontinuierlicher, zyklischer Prozess, der über mehrere Jahre wirkt.

Eine BI-Strategie wird normalerweise in mehreren Phasen entwickelt und dauert normalerweise etwa drei bis sechs Monate. Es folgt eine ständige Verbesserung der Komponenten, wie beispielsweise der Planung der Maßnahmen, durch das Anforderungsmanagement, das entweder monatlich oder jährlich durch das Budgetierungsverfahren durchgeführt wird. Die BI-Strategie sollte regelmäßig überprüft und an die Ziele des Unternehmens angepasst werden. ⁴²

Im Anschluss soll auch der Ablauf der Entwicklung einer BI-Strategie ähnlich dem strategischen Entwicklungsprozess des Unternehmens definiert werden. Die folgenden Phasen der eigentlichen Umsetzung sowie das Controlling der Umsetzung sind nicht Teil der Entwicklung und werden daher nicht in der Abbildung 5 dargestellt. ⁴³

⁴¹ (vgl. Gansor et al. 2010, S. 39)

⁴² (vgl. ebd.: S. 73-75)

⁴³ (vgl. ebd.: 2010 S. 75)

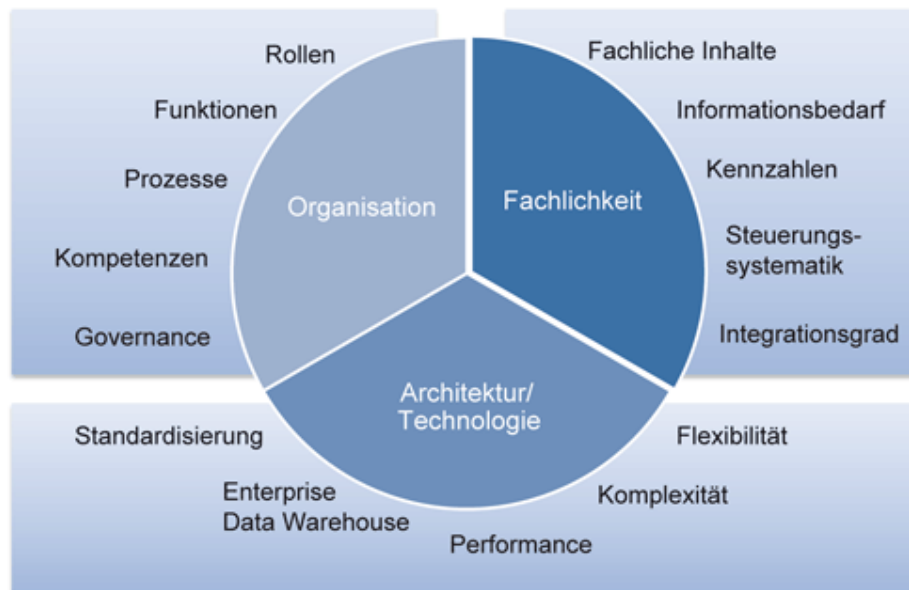


Abbildung 5: Perspektiven einer BI-Strategie
(Quelle: vgl. Chamoni/Glichowski 2004, S.120)

Es ist grundsätzlich möglich, dass strategische Projekte im Bereich der Informationsverarbeitung nur eine der drei Perspektiven anwenden. Eine holistische (ganzheitliche) Methode zur Entwicklung einer BI-Strategie umfasst allerdings die Berücksichtigung aller drei Standpunkte.

Bei der Entwicklung einer spezifischen Strategie können die Schwerpunkte je nach Perspektive variieren, je nachdem, welche Priorität die Anforderungen aus den drei Bereichen für das jeweilige Unternehmen haben und auf welcher bereits vorhandenen Basis aufgesetzt werden können.⁴⁴

Wenn der Phasenablauf der Strategieentwicklung mit den Perspektiven kombiniert wird, entsteht eine matrixartige Struktur, die sich in zehn Aktivitätenblöcke unterteilt, wie in Abbildung 6.

Es ist wichtig, die Darstellung nicht so zu verstehen, dass alle drei Standpunkte gleichzeitig abgearbeitet werden müssen. Es besteht die Möglichkeit, die Standpunkte entweder parallel (Fachlichkeit - Architektur - Technologie - Organisation) oder in Kombination (Analyse, Bewertung, Konzeption und Maßnahmen) zu entwickeln. Die Untersuchung der Fachlichkeit ist sinnvollerweise vor Architektur und Technologie. Oftmals sind auch informatorische und architektonische Aspekte vor der Formulierung von Aussagen zur Organisation zu planen.⁴⁵

⁴⁴ (vgl. Gansor et al. 2010 S. 74)

⁴⁵ (vgl. ebd.: S. 75)

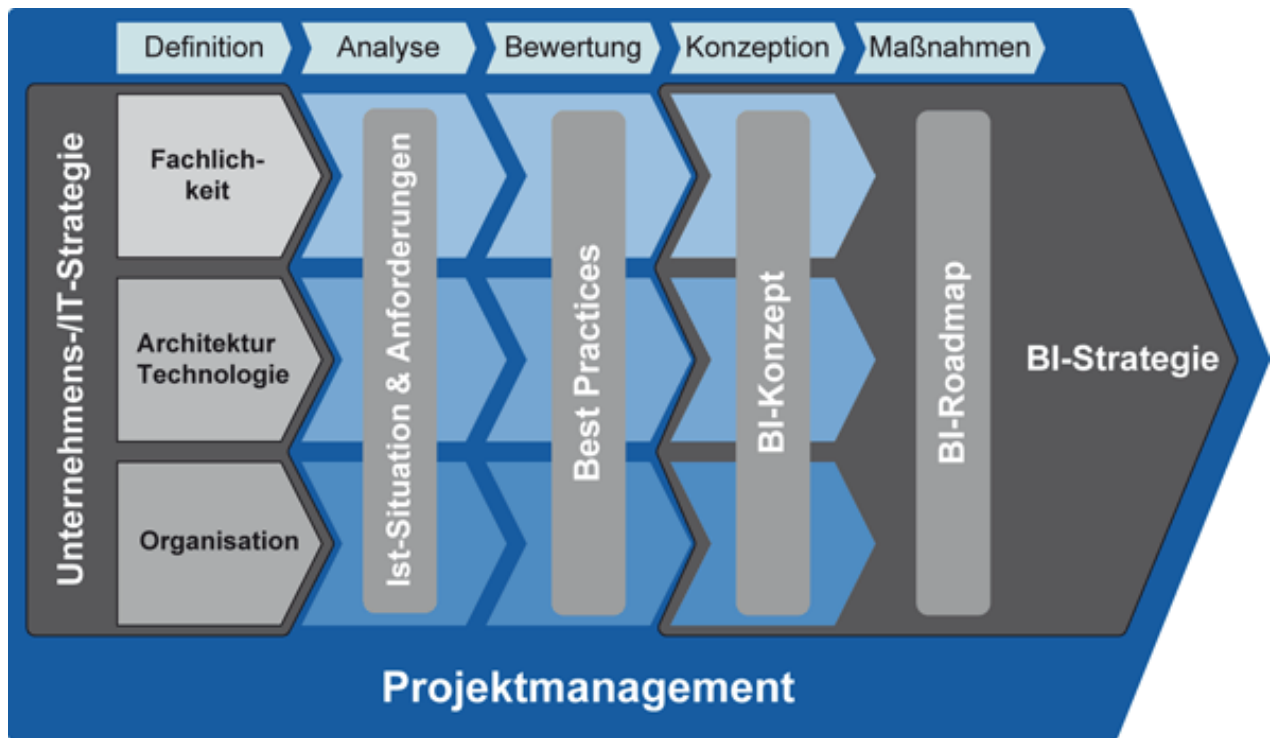


Abbildung 6 : Ganzheitliches Vorgehensmodell zur Entwicklung der BI Strategie

(Quelle : Von der Strategie zum BICC)

Eine Business Intelligence-Strategie wird in der Regel unter Berücksichtigung verschiedener Perspektiven entwickelt. Wie in Abbildung 5 veranschaulicht, spielen dabei drei zentrale Blickwinkel eine Rolle: Fachlichkeit, Architektur/Technologie und Organisation. Es ist durchaus möglich, dass strategische IT-Projekte sich ausschließlich auf eine dieser Perspektiven konzentrieren. Zum Beispiel kann die Ausrichtung auf die Fachlichkeit im Vordergrund stehen, wenn ein Unternehmen ein neues Steuerungssystem wie die Balanced Scorecard entwickelt. Dennoch ist eine umfassende Methode zur Entwicklung einer BI-Strategie darauf ausgerichtet, alle drei Standpunkte zu berücksichtigen.⁴⁶

Während der Ausarbeitung einer spezifischen Strategie können die Schwerpunkte je nach Perspektive variieren. Dies hängt davon ab, welche Anforderungen in den drei Bereichen für das jeweilige Unternehmen von höchster Bedeutung sind und auf welcher vorhandenen Grundlage aufgebaut werden kann. Eine holistische Herangehensweise gewährleistet, dass die BI-Strategie den Bedürfnissen des Unternehmens gerecht wird und effektive Geschäftslösungen bereitstellt.⁴⁷

Der Phasenablauf der Strategieentwicklung wird in Abbildung 6, wie von Tom Gansor et al. beschrieben, mit den verschiedenen Perspektiven kombiniert. Dabei ergibt sich eine matrixartige Struktur, die sich in zehn Aktivitätenblöcke unterteilen lässt.

⁴⁶ (vgl. Gansor et al. 2010 S. 74)

⁴⁷ (vgl. ebd.: S. 75)

Die Phasen definieren die Reihenfolge der Entwicklungsschritte, während die Perspektiven die inhaltliche Ausrichtung auf die zu betrachtenden Themen bestimmen.

Im Laufe des Entwicklungsprozesses werden gemäß Tom Gansor et al. verschiedene Deliverables bereitgestellt. Diese sind in der Regel wesentlich für die strategische Ausrichtung des Unternehmens. Es ist wichtig, diese Vorgehensweise flexibel an die spezifischen Anforderungen und Gegebenheiten anzupassen, um die bestmöglichen Ergebnisse zu erzielen.

Gemäß den Erklärungen von Tom Gansor et al. beinhaltet die Entwicklung einer Business Intelligence-Strategie eine Reihe von wichtigen Schritten, die es zu berücksichtigen gilt. Diese Schritte sind entscheidend, um sicherzustellen, dass die BI-Initiative erfolgreich ist und den Bedürfnissen des Unternehmens gerecht wird.

Zunächst einmal steht die Ist-Dokumentation und die Erhebung der Anforderungen im Mittelpunkt. Hierbei handelt es sich um eine gründliche Analyse der aktuellen Situation und um die Identifizierung der Anforderungen, die erfüllt werden müssen. Dies kann beispielsweise durch eine Informationsbedarfsanalyse erfolgen.

Ein weiterer wichtiger Schritt ist die Bewertung des BI-Reifegrads des Unternehmens. Hierbei wird die Ist-Situation im Vergleich zu den Anforderungen und anderen Unternehmen analysiert. Methoden wie Gap-/SWOT-Analysen und Benchmarking sind nützlich, um diese Bewertung vorzunehmen.

Das BI-Konzept umfasst die Definition der Vision, der Ziele, des Informationsangebots, der Architektur und des Anwendungsportfolios. Dies ist von grundlegender Bedeutung, um die strategische Ausrichtung der BI-Initiative festzulegen.

Schließlich ist die Entwicklung einer BI-Roadmap entscheidend, um einen schrittweisen Maßnahmenplan und ein Projektportfolio zur schrittweisen Umsetzung des Konzepts zu erstellen.

Diese Schritte sind unverzichtbar für die Entwicklung einer effektiven BI-Strategie und bilden die Grundlage für den Erfolg in diesem Bereich. Wie von Tom Gansor et al. erläutert, sollte dieser Prozess sorgfältig geplant und angepasst werden, um die spezifischen Anforderungen und Gegebenheiten des Unternehmens zu berücksichtigen.

Kapitel 3 : Data Governance

In diesem Abschnitt wird die Bedeutung der Datenschutzthematik im Geschäftsumfeld behandelt. Hierbei werden spezifische Vorgaben und Regelungen seitens des Gesetzgebers präsentiert und eingehend erörtert.

3.1. Data Governance und Datenethik

In der heutigen datengetriebenen Welt gewinnt Data Governance zunehmend an Bedeutung. Sie ist der Schlüssel zur Gewährleistung, dass Daten in jeder Phase ihres Lebenszyklus in Übereinstimmung mit den geltenden Gesetzen und Vorschriften erfasst, gespeichert, aufbewahrt und genutzt werden. Data Governance umfasst weit mehr als nur das Festlegen von Richtlinien und Verfahren – sie bildet das Fundament für die datenbasierte Entscheidungsfindung, Datenschutz und Compliance in einer Organisation.⁴⁸

Deutschland ist seit jeher ein Vorreiter bei der Entwicklung des Datenschutzrechts. Das weltweit erste formelle Datenschutzgesetz wurde 1970 vom Bundesland Hessen erlassen. Ein Bundesdatenschutzgesetz folgte im Jahr 1977. Seit 1995, mit der Europäischen Datenschutzrichtlinie ist das Datenschutzrecht in Deutschland durch einen europäischen Rahmen geprägt, der die Grundlagen des Datenschutzrechts in allen Mitgliedstaaten der Europäischen Gemeinschaften/Union bestimmt (vgl. Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr. o. D.).

Diese europäische Richtlinie wurde durch die 2018 in Kraft getretene "General Data Protection Regulation (GDPR)" abgelöst.⁴⁹ Während eine Richtlinie ein Rahmenwerk ist, das den Inhalt des nationalen Rechts prägt und den nationalen Gesetzgebern einen Handlungsrahmen vorgibt, ist eine Verordnung unmittelbar anwendbar, ohne dass ein Durchführungsgesetz erlassen werden muss.

Im Gegensatz zu anderen Verordnungen lässt die DSGVO den Mitgliedstaaten jedoch immer noch einen erheblichen Spielraum zur Regulierung (was sie gewissermaßen zu einer "Mischform" zwischen einer Verordnung und einer Richtlinie macht), was zu einer Zusammenarbeit zwischen der DSGVO und einem neuen Bundesdatenschutzgesetz sowie anderen nationalen Datenschutzgesetzen geführt hat, die auch dann noch als Durchführungsrechtsakte dienen, wenn die DSGVO eine Verordnung ist.⁵⁰

In ihrer aktuellen Fassung lauten diese Grundsätze: Verbot der Verarbeitung ohne Erlaubnis = Rechtmäßigkeit der Verarbeitung; Fairness; Transparenz; Zweckbindung; Datenminimierung; Richtigkeit; Speicherbegrenzung; Integrität und Vertraulichkeit; Verantwortlichkeit.

⁴⁸ (vgl. Poßner 2023, S. 7)

⁴⁹ (vgl. Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG. 2016)

⁵⁰ (vgl. Poßner 2023, S. 7)

Zu bedenken ist auch, dass das Ziel des Datenschutzrechts der Schutz des Rechts des Einzelnen auf informationelle Selbstbestimmung ist; der Schutz der Daten ist nur ein Mittel zu diesem Zweck.

Eine generelle und unabdingbare Voraussetzung für die Verarbeitung personenbezogener Daten ist es, diese Daten auf einem technischen und organisatorischen Niveau zu sichern, das den aus der Verarbeitung resultierenden Risiken angemessen ist. Die Umsetzung der Datensicherheit ("Sicherheit der Verarbeitung") ist in Artikel 32 DSGVO festgelegt. Damit gehen Datenschutz und Daten-/IT-Sicherheit Hand in Hand.⁵¹

Als BI-Experte ist man dafür verantwortlich, Daten ethisch zu behandeln. Datenethik bezieht sich auf fundierte Standards für richtig und falsch, die vorschreiben, wie Daten gesammelt, weitergegeben und verwendet werden. Dazu gehören in der Regel personenbezogene Daten, die allein oder zusammen mit anderen Daten verwendet werden können, um die Identität einer Person festzustellen. Ein Element des ethischen Umgangs mit diesen Daten ist die Gewährleistung des Datenschutzes und der Sicherheit dieser Daten während ihrer gesamten Lebensdauer.

3.2. Privatsphäre ist wichtig

Datenschutz bedeutet, dass die Informationen und Aktivitäten einer betroffenen Person bei jeder Datentransaktion geschützt werden. Dies wird auch als Informationsschutz oder Datenschutz bezeichnet. Der Datenschutz befasst sich mit dem Zugang zu, der Verwendung und der Sammlung von personenbezogenen Daten. Für die Personen, deren Daten gesammelt werden, bedeutet dies, dass sie laut Datenschutzgrundverordnungen ein Recht darauf haben:

- Freiheit von der unangemessenen Verwendung ihrer Daten
- das Recht auf Einsichtnahme, Aktualisierung oder Berichtigung ihrer Daten
- die Möglichkeit, der Datenerhebung zuzustimmen
- Rechtsanspruch auf Zugang zu den Daten

Um diese Rechte zu wahren, müssen Unternehmen und Organisationen Maßnahmen zum Schutz der Privatsphäre ergreifen, um die Daten des Einzelnen zu schützen. Dies ist auch eine Frage des Vertrauens.

Es ist wichtig, dass die Öffentlichkeit den Unternehmen im Umgang mit persönlichen Daten vertraut. Das ist der Grund, warum Menschen die Produkte eines Unternehmens nutzen, ihre Daten weitergeben und vieles mehr. Vertrauen ist eine große Verantwortung, die man nicht auf die leichte Schulter nehmen darf.

Unternehmen müssen viele verschiedene Maßnahmen einsetzen, um die Privatsphäre der betroffenen Personen zu schützen, z. B. durch die Vergabe von Zugriffsrechten, um sicherzustellen, dass nur die Personen Zugriff auf die Daten haben, die dazu berechtigt sind. Eine weitere wichtige Strategie zur Wahrung der Privatsphäre ist die Anonymisierung von Daten.

⁵¹ (vgl. Poßner 2023, S. 6)

Bei der Datenanonymisierung werden die privaten oder sensiblen Daten von Personen durch Eliminierung personenbezogener Daten geschützt. In der Regel beinhaltet die Datenanonymisierung das Ausblenden, Verschlüsseln oder Maskieren persönlicher Informationen, oft durch die Verwendung von Codes mit fester Länge zur Darstellung von Datenspalten oder das Ausblenden von Daten mit veränderten Werten.

Die Anonymisierung von Daten wird in nahezu jeder Branche eingesetzt. Als BI-Experte wird man wahrscheinlich nicht persönlich mit der Anonymisierung arbeiten, aber es ist nützlich zu wissen, welche Arten von Daten häufig anonymisiert werden, bevor man damit arbeitet. Zu diesen Daten können gehören:

- Telefonnummern
- Namen
- Kfz-Kennzeichen und Nummernschilder
- Sozialversicherungsnummern
- IP-Adressen
- Medizinische Daten
- E-Mail-Adressen
- Fotografien
- Kontonummern

Kapitel 4 : Anforderungsanalyse & Praktische Umsetzung

Folgendes Kapitel widmet sich der praktischen Implementierung des Business Intelligence-Ansatzes. Hierbei erfolgt zunächst eine Ist-Analyse, gefolgt von der gezielten Abfrage spezifischer Daten aus der Datenbank. Diese werden daraufhin mittels Visualisierungswerkzeugen in unterschiedlichen Diagrammen veranschaulicht.

5.1. Anforderungsanalyse und Projektplanung

4.1.1 Projektziel: Implementierung eines effektiven Business-Intelligence-Ansatzes, um die Nutzung der Fahrrädern zu verstehen und den Kundenstamm zu erweitern.

4.1.2 Aufgaben von der Abteilung für Kundenentwicklung:

Verstehen, was Kunden wollen, was ein erfolgreiches Produkt ausmacht und wie neue Stationen die Nachfrage in verschiedenen geografischen Gebieten lindern könnten.

Verstehen, wie das aktuelle Fahrradsortiment genutzt wird.

Wie kann man die Erkenntnisse über die Kundennutzung nutzen, um neue Stationen zu entwickeln?

Wie verschiedene Nutzer (Abonnenten und Nicht-Abonnenten) die Fahrräder benutzen. Es wird dabei eine große Gruppe von Nutzern untersucht, um eine Repräsentation der Nutzer an verschiedenen Standorten und mit niedrigen bis hohen Aktivitätsniveaus zu erhalten.

Wie beeinflusst das Wetter die Nutzung von Fahrrädern?

4.1.3 Ergebnisse und Metriken:

Eine graphische oder kartografische Darstellung, in der die Standorte der Start- und Zielstationen nach Orten zusammengefasst sind. Man kann dabei jede beliebige Ortsbezeichnung verwenden, z. B. Name der Station, Postleitzahl, Stadtviertel und/oder Bezirk. Dies sollte die Anzahl der Fahrten an den Startorten zeigen.

Visualisierungen, die zeigen, welche Zielstandorte (Endstandorte) auf der Grundlage der gesamten Fahrminuten beliebt sind.

Visualisierungen, die sich auf die Trends im Sommer 2022 konzentrieren.

Visualisierungen, die das Wachstum der Anzahl der Fahrten von Jahr zu Jahr zeigt.

Erkenntnisse über die Anzahl der Fahrten an allen Start- und Endstationen.

Erkenntnisse über die Nutzungsspitzen nach Monats- und Jahreszeit sowie über die Auswirkungen des Wetters.

4.1.4 Erfolgskriterien:

Die Erfolgskriterien für den Business Intelligence (BI)-Ansatz sind von entscheidender Bedeutung, um sicherzustellen, dass der Geschäftsbetrieb effizient und effektiv gesteuert wird. Die korrekte Abfrage von Daten steht an vorderster Stelle, da sie die Grundlage für fundierte Entscheidungen bildet.

Sauberkeit der Daten ist ein weiterer entscheidender Faktor. Es muss dabei sichergestellt werden, dass die Datenquellen frei von Unreinheiten und Fehlern sind, da dies die Qualität der Analysen beeinflusst.

Darüber hinaus ist die klare Sichtbarkeit der Ergebnisse ein Muss. Die BI-Analysen sollen verständlich und leicht interpretierbar sein, damit sie den Stakeholdern und den Entscheidungsträgern bei der Identifizierung von Trends und der Beantwortung geschäftlicher Fragen helfen können.

Schließlich ist es das Ziel, nicht nur Daten zu analysieren, sondern auch Antworten auf geschäftliche Fragen zu liefern. Um sicherzustellen, dass die BI-Analysen einen Mehrwert für das Unternehmen bieten und konkrete Handlungsempfehlungen liefern.

5.2. Praktische Umsetzung

4.2.1 Datenquellen

- i) Global Surface Summary of the Day Weather Data : Dieser öffentliche Datensatz wurde von der National Oceanic and Atmospheric Administration (NOAA) erstellt und enthält globale Daten, die vom USAF Climatology Center stammen. Dieser Datensatz umfasst GSOD-Daten zwischen 1929 und heute (täglich aktualisiert), die von über 9000 Stationen gesammelt wurden.

Die globale Zusammenfassung des Tages besteht aus einem Dutzend Tagesmittelwerten, die aus den stündlichen Daten der weltweiten Stationen berechnet werden. Zu den täglichen Wetterelementen gehören Mittelwerte für: Temperatur, Taupunkttemperatur, Luftdruck auf Meereshöhe, Stationsdruck, Sichtweite und Windgeschwindigkeit sowie Höchst- und Mindesttemperatur, maximale anhaltende Windgeschwindigkeit und maximale Windböen, Niederschlagsmenge, Schneehöhe und Wetterindikatoren. Mit Ausnahme der US-Stationen basieren die 24-Stunden-Zeiträume auf der UTC-Zeit.

- ii) Census Bureau US Boundaries : Es handelt sich in diesem Dataset um Grenzdateien in voller Auflösung, die von TIGER/Line Shapefiles, den vollständig unterstützten geografischen Kernprodukten des US Census Bureau, abgeleitet sind. Es handelt sich um Auszüge ausgewählter geografischer und kartografischer Informationen aus der Datenbank Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) des US Census Bureau.

Dazu gehören Informationen für die 50 Bundesstaaten, den District of Columbia, Puerto Rico und die Inselgebiete (Amerikanisch-Samoa, das Commonwealth der Nördlichen Marianen, Guam und die Jungferninseln der Vereinigten Staaten). Dazu gehören Polygongrenzen von geografischen und statistischen Gebieten, lineare Merkmale, einschließlich Straßen und Hydrographie, und Punktmerkmale.

Der Datensatz beschreibt die Grenzen der folgenden US-Gebiete:

- (1) Staaten
- (2) Landkreise
- (3) 115. und 116. US-Kongressbezirke
- (4) Städtische und mikropolitane statistische Gebiete
- (5) Städtische Gebiete
- (6) Stadtgrenzen
- (7) Kombinierte statistische Gebiete
- (8) Küstenlinie
- (9) Besiedelte Orte
- (10) Nationale Grenze

iii) NYC Citi Bike Trips : Citi Bike ist mit 10.000 Fahrrädern und 600 Stationen in Manhattan, Brooklyn, Queens und Jersey City das größte Fahrradverleihprogramm der USA. Dieser Datensatz enthält Citi Bike-Fahrten seit dem Start von Citi Bike im September 2013 und wird täglich aktualisiert. Die Daten wurden von Citi Bike verarbeitet, um Fahrten zu entfernen, die von Mitarbeitern zur Wartung und Inspektion des Systems unternommen werden, sowie Fahrten von weniger als 60 Sekunden Dauer, die als Fehlstarts gelten.

iv) NYC Zip Codes : Eine eigenständige Google Sheets-Datei mit den Postleitzahlen (PLZ) für die Stadt New York, die dazu dient, geografische Informationen zu organisieren und darzustellen.

4.2.2 Datensammlung, Datenmodellierung und ETL-Prozesse

4.2.2.1 Arten von Schemata (Stern und Schneeflocke)

Stern- und Schneeflockenschemata haben einige Gemeinsamkeiten, aber auch ein paar Unterschiede. Obwohl beide Schemata Dimensionstabellen verwenden, sind die Dimensionstabellen in Schneeflockenschemata normalisiert. Dadurch werden die Daten in zusätzliche Tabellen aufgeteilt, was die Schemata ein wenig komplexer macht.

Ein Sternschema ist ein Schema, das aus einer oder mehreren Faktentabellen besteht, die auf eine beliebige Anzahl von Dimensionstabellen verweisen.

Wie der Name schon sagt, hat dieses Schema die Form eines Sterns. Diese Art von Schema ist ideal für die Bereitstellung von Informationen in großem Umfang und macht die Leseausgabe effizienter.

Im Folgenden werden zwei Schneeflockenschemata präsentiert, die speziell für diese Arbeit von Relevanz sind. Sie werden als Schneeflockenschemata bezeichnet, da sie etwas komplexer strukturiert sind als Sternschemata. Beide Schemata enthalten eine Zieltabelle, in die Daten aus verschiedenen relevanten Quellen einfließen. Diese bilden die Grundlage für die Datenanalyse und Datenvisualisierung. Die Schemata wurden entwickelt, um einen Überblick über die Quellen zu behalten, damit bei Bedarf darauf zurückgegriffen werden kann.



Abbildung 7: Schneeflockenschema-Datenmodell für erste Zieltabelle

(Quelle: Eigene Darstellung mittels dbdiagramm.io)

In der Abbildung 7 wird das Datenmodell für die erste Zieltabelle präsentiert. Diese enthält alle Geschäftsdaten, die von 2018 bis 2023 erfasst wurden.



Abbildung 8: Schneeflockenschema-Datenmodell für zweite Zieltabelle
(Quelle: Eigene Darstellung mittels dbdiagramm.io)

In der Abbildung 8 wird das Datenmodell für die zweite Zieltabelle präsentiert. Diese enthält Geschäftsdaten für die Sommersaison 2022.

4.2.3 Google Cloud Plattform + Tableau

4.2.3.1 Google Data Flow

Google Dataflow bietet eine Vielzahl von leistungsstarken Funktionen für die Datenverarbeitung, darunter die nahtlose Synchronisation oder Replikation von Daten über verschiedene Datenquellen hinweg. Die effiziente Integration und Übertragung von Daten in verschiedenen Umgebungen werden durch diese Funktion gewährleistet.

Darüber hinaus haben Google Dataflow intelligente Diagnosefunktionen, die Pipeline-Probleme identifizieren können. Diese hochmodernen Diagnosewerkzeuge ermöglichen eine proaktive Identifizierung von Problemen in Datenpipelines, um einen reibungslosen Ablauf sicherzustellen und mögliche Engpässe frühzeitig zu erkennen.⁵²

Eine weitere Funktion ist die Verwendung von SQL, um Daten-Pipelines anhand der Benutzeroberfläche von BigQuery zu erstellen. Mithilfe einer intuitiven Schnittstelle ist es möglich, anspruchsvolle Datenverarbeitungsabläufe effektiv zu gestalten.

Die Möglichkeit, Ressourcen zu planen, um die Kosten für die Stapelverarbeitung zu optimieren, ist ein wichtiger Aspekt von Google Dataflow. Dies umfasst die Entwicklung und Umsetzung von Methoden zur effektiven Nutzung von Ressourcen, um die Gesamtkosten für die Durchführung von Datenverarbeitungsaufgaben zu senken.⁵³

Die Verwendung von Pipeline-Vorlagen erleichtert den Start des Pipeline-Erstellungsprozesses und fördert die gemeinsame Nutzung von Systemen in einem Unternehmen. Teams können mit vorgefertigten Vorlagen effektiver zusammenarbeiten und bewährte Datenverarbeitungspraktiken gemeinsam nutzen.

Insgesamt stellt Google Dataflow mit seinen vielfältigen Funktionen eine leistungsfähige Plattform für die Datenverarbeitung dar. Es zeichnet sich durch eine nahtlose Integration, intelligente Diagnosemöglichkeiten, benutzerfreundliche SQL-Entwicklung, Kostenoptimierung und die Nutzung von Vorlagen für eine effektive Zusammenarbeit in Unternehmen aus.⁵⁴

⁵² (vgl. Dataflow – Übersicht o. D.)

⁵³ (vgl. ebd.)

⁵⁴ (vgl. ebd.)

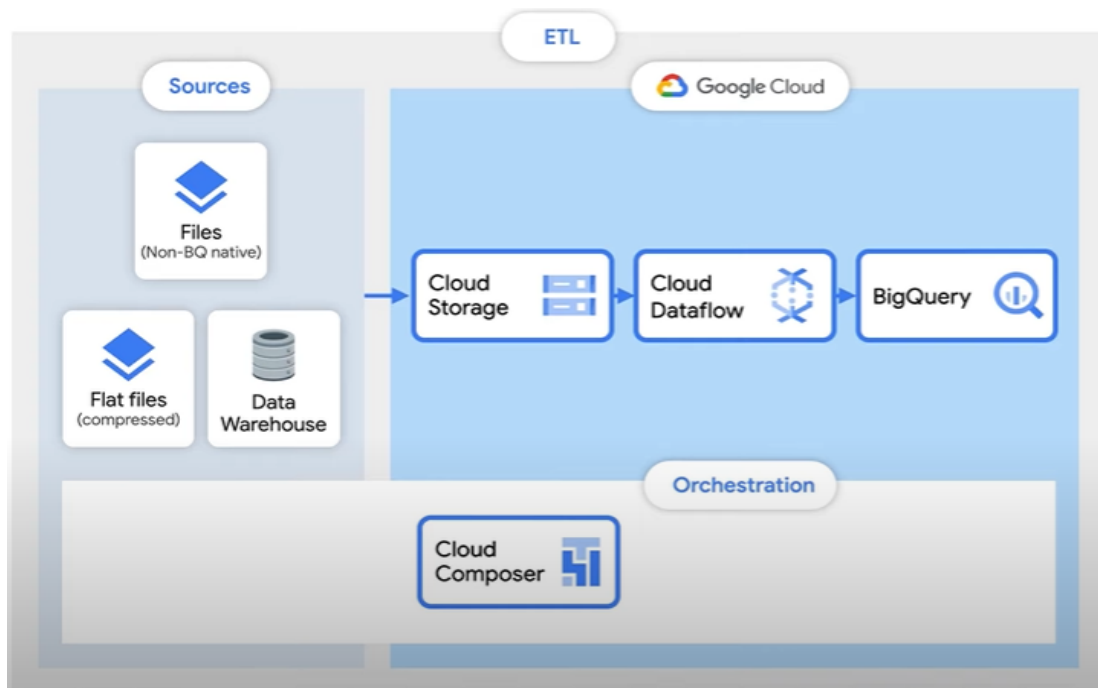


Abbildung 9: Google Cloud Platform Architektur – ETL

(Quelle: Eigene Darstellung)

4.2.3.2 Google Big-Query

BigQuery ist ein vollständig verwaltetes Data Warehouse für Unternehmen, das maschinelles Lernen, raumbezogene Analysen und Business Intelligence integriert, um die Daten zu verwalten und zu analysieren. Dank der serverlosen Architektur von BigQuery kann man SQL-Abfragen verwenden, um die wichtigsten Probleme jeder Organisation ohne Infrastrukturverwaltung zu lösen. Die skalierbare, verteilte Analyse-Engine von BigQuery ermöglicht die Abfrage von Terabyte in Sekunden und Petabyte in Minuten.⁵⁵

Durch die Trennung der Rechen-Engine, die die Daten von Ihren Speicheroptionen analysiert, erhöht BigQuery die Flexibilität. BigQuery bietet die Möglichkeit, Ihre Daten zu speichern und zu analysieren oder zu bewerten. Föderale Abfragen ermöglichen das Lesen von Daten aus externen Quellen, während Streaming kontinuierliche Datenaktualisierungen unterstützt.⁵⁶

Darüber hinaus ist das Interface benutzerfreundlich und angenehm, was zur einfachen Nutzung der Plattform beiträgt.

⁵⁵ (vgl. Was ist BigQuery? o. D.)

⁵⁶ (vgl. ebd.)

4.2.3.3 SQL-Abfrage zur Erstellung von der ersten Ziel-Tabelle (für den ganzen Zeitraum)

```
SELECT
  NYB.usertype,
  ZIPSTART.zip_code AS zip_code_start,
  ZIPSTARTNAME.borough AS borough_start,
  ZIPSTARTNAME.neighborhood AS neighborhood_start,
  ZIPEND.zip_code AS zip_code_end,
  ZIPENDNAME.borough AS borough_end,
  ZIPENDNAME.neighborhood AS neighborhood_end,
  DATE_ADD(DATE(NYB.starttime), INTERVAL 5 YEAR) AS start_day,
  DATE_ADD(DATE(NYB.stoptime), INTERVAL 5 YEAR) AS stop_day,
  WEA.temp/2 AS day_mean_temperature, -- Durchschnittliche Temperatur
  WEA.wdsp AS day_mean_wind_speed, -- Durchschnittliche Windgeschwindigkeit
  WEA.prcp AS day_total_precipitation, -- Totale
  Niederschlagswahrscheinlichkeit
  -- Group trips into 10 minute intervals to reduces the number of rows
  ROUND(CAST(NYB.tripduration / 60 AS INT64), -1) AS trip_minutes,
  COUNT(NYB.bikeid) AS trip_count

FROM
  `bigquery-public-data.new_york_citibike.citibike_trips` AS NYB
  INNER JOIN
    `bigquery-public-data.geo_us_boundaries.zip_codes` ZIPSTART
  ON ST_WITHIN(
    ST_GEOGPOINT(NYB.start_station_longitude, NYB.start_station_latitude),
    ZIPSTART.zip_code_geom)
  INNER JOIN
    `bigquery-public-data.geo_us_boundaries.zip_codes` ZIPEND
  ON ST_WITHIN(
    ST_GEOGPOINT(NYB.end_station_longitude, NYB.end_station_latitude),
    ZIPEND.zip_code_geom)
  INNER JOIN
    `bigquery-public-data.noaa_gsod.gsod20*` AS WEA
  ON PARSE_DATE("%Y%m%d", CONCAT(WEA.year, WEA.mo, WEA.da)) =
    DATE(NYB.starttime)
  INNER JOIN
    -- Note! zip code table name, enclosed in backticks
    `vrfvrvrvrvrvrfr.Zip_Codes.NYC_zip_codes` AS ZIPSTARTNAME
  ON ZIPSTART.zip_code = CAST(ZIPSTARTNAME.zip AS STRING)
  INNER JOIN
    -- Note! zipcode table name, enclosed in backticks
    `vrfvrvrvrvrvrfr.Zip_Codes.NYC_zip_codes` AS ZIPENDNAME
  ON ZIPEND.zip_code = CAST(ZIPENDNAME.zip AS STRING)
WHERE
  -- This takes the weather data from one weather station
  WEA.wban = '94728' -- NEW YORK CENTRAL PARK

GROUP BY
  1,
  2,
  3,
  4,
  5,
  6,
  7,
```

8,
9,
10,
11,
12,
13

Nach Ausführung des Codes in Google BigQuery entsteht eine Ziel-Tabelle, die Daten für den Zeitraum (2018 bis 2023) repräsentiert. Die Tabelle besteht aus 6.492.316 Einträgen und ist durch die Integration mehrerer Datenquellen erstellt. Die Resultate der SQL-Abfrage wurden in Form einer CSV-Datei auf Google Drive gespeichert.

4.2.3.4 SQL-Abfrage zur Erstellung von Ziel-Tabelle (für das Sommer -Saison)

```
SELECT
TRI.usertype,
TRI.start_station_longitude,
TRI.start_station_latitude,
TRI.end_station_longitude,
TRI.end_station_latitude,
ZIPSTART.zip_code AS zip_code_start,
ZIPSTARTNAME.borough AS borough_start,
ZIPSTARTNAME.neighborhood AS neighborhood_start,
ZIPEND.zip_code AS zip_code_end,
ZIPENDNAME.borough AS borough_end,
ZIPENDNAME.neighborhood AS neighborhood_end,
-- Da die Fahrten aus den Jahren 2014 und 2015 stammen, es werden 7 Jahre
-- hinzugefügt, damit das Datum aktuell aussieht.
DATE_ADD(DATE(TRI.starttime), INTERVAL 7 YEAR) AS start_day,
DATE_ADD(DATE(TRI.stoptime), INTERVAL 7 YEAR) AS stop_day,
WEA.temp / 2 AS day_mean_temperature, -- durchschnittliche temp
WEA.wdsp AS day_mean_wind_speed, -- durchschnittliche windgeschwindigkeit
WEA.prcp AS day_total_precipitation, -- Total niederschlagswahrscheinlichkeit
-- Wir werden die Fahrten in 10-Minuten-Intervallen gruppieren, was auch die
-- Anzahl der Zeilen reduziert.
ROUND(CAST(TRI.tripduration / 60 AS INT64), -1) AS trip_minutes,
TRI.bikeid
FROM
`bigquery-public-data.new_york_citibike.citibike_trips` AS TRI
INNER JOIN
`bigquery-public-data.geo_us_boundaries.zip_codes` ZIPSTART
ON ST_WITHIN(
ST_GEOGPOINT(TRI.start_station_longitude, TRI.start_station_latitude),
ZIPSTART.zip_code_geom)
INNER JOIN
`bigquery-public-data.geo_us_boundaries.zip_codes` ZIPEND
ON ST_WITHIN(
ST_GEOGPOINT(TRI.end_station_longitude, TRI.end_station_latitude),
ZIPEND.zip_code_geom)
INNER JOIN
-- https://pantheon.corp.google.com/bigquery?p=bigquery-public-
-- data&d=noaa_gsod
`bigquery-public-data.noaa_gsod.gsod20*` AS WEA
ON PARSE_DATE("%Y%m%d", CONCAT(WEA.year, WEA.mo, WEA.da)) =
DATE(TRI.starttime)
INNER JOIN
-- Name der Postleitzahlentabelle
`vrfvrvrvrvrvrfr.Zip_Codes.NYC_zip_codes` AS ZIPSTARTNAME
ON ZIPSTART.zip_code = CAST(ZIPSTARTNAME.zip AS STRING)
INNER JOIN
`vrfvrvrvrvrvrfr.Zip_Codes.NYC_zip_codes` AS ZIPENDNAME
ON ZIPEND.zip_code = CAST(ZIPENDNAME.zip AS STRING)
WHERE
-- das Wetter von einer Wetterstation "NYC", diese deckt die Stadt New York
-- ab
WEA.wban = '94728' -- NEW YORK CENTRAL PARK
-- Daten für drei Sommermonate
AND DATE(TRI.starttime) BETWEEN DATE('2015-07-01') AND DATE('2015-09-30')
```

Nach Ausführung des SQL-Codes in Google BigQuery entsteht eine Ziel-Tabelle, die Daten für das gesamte Sommer-Saison repräsentiert. Die Tabelle besteht aus 3.499.028 Einträgen und ist durch die Integration mehrerer Datenquellen erstellt. Die Resultate der SQL-Abfrage wurden in Form einer CSV-Datei auf Google Drive gespeichert.

4.2.3.5 Datenvisualisierung & Gewinnung von Erkenntnissen

Nach der Generierung von Ziel-Tabellen mithilfe von SQL-Abfragen in Google Cloud Plattform (GCP) wurden die resultierenden Daten in Google Docs exportiert. Anschließend wurde eine Verbindung zu Tableau Software hergestellt, um eine eingehende Analyse und die Erstellung von aussagekräftigen Grafiken zu ermöglichen. Dieser Prozess erlaubte es, die umfangreiche Datenmenge auf eine ansprechende und leicht verständliche Weise zu visualisieren, wodurch tiefergehende Erkenntnisse gewonnen werden konnten.

Tableau bietet leistungsstarke Funktionen zur Datenvisualisierung, die es ermöglichen, komplexe Datenmuster, Trends und Beziehungen auf einfache Weise zu identifizieren. Die erstellten Grafiken bieten nicht nur eine klare Darstellung der Daten, sondern ermöglichen auch eine interaktive Exploration. Benutzer können durch die Daten navigieren, Filter anwenden und spezifische Aspekte der Daten näher betrachten.

Die Visualisierung der Daten in Tableau trägt dazu bei, wichtige Erkenntnisse zu gewinnen, indem sie komplexe Zusammenhänge in den Daten verdeutlicht. Diese visuelle Herangehensweise unterstützt effektiv die Entscheidungsfindung und hilft, Muster zu erkennen, die in reinen Tabellendaten möglicherweise übersehen worden wären. Zusammenfassend erlaubt diese Integration von SQL-Abfragen, Google Docs und Tableau eine umfassende Datenanalyse und fördert das Verständnis von relevanten Informationen für datengetriebene Entscheidungsprozesse.

4.2.3.5.1 Datenvisualisierung, /-Analyse für den Zeitraum (2018-2022)

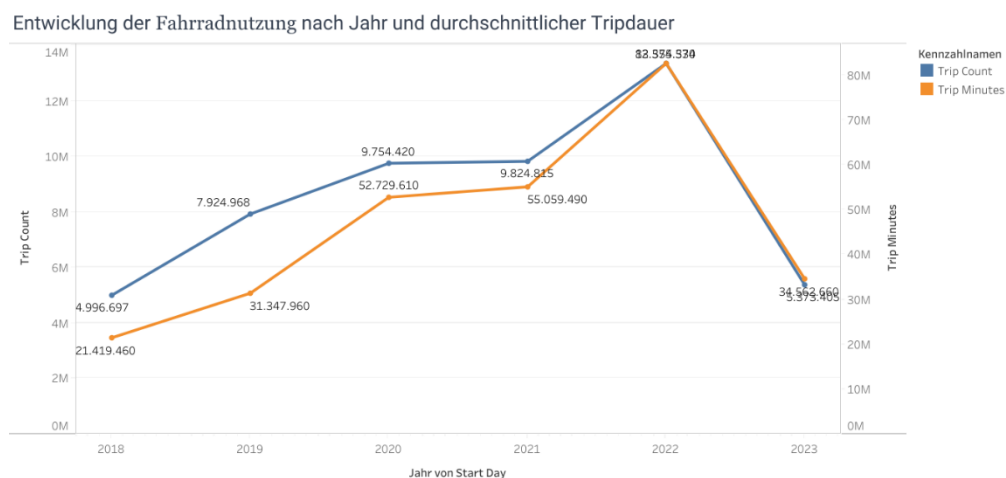


Abbildung 12 : Jahresvergleich: Fahrradnutzung und durchschnittliche Tripdauer

(Quelle : Tableau Software – eigene Darstellung)

Die visuelle Darstellung zeigt die Entwicklung der Fahrradnutzung von 2018 bis 2023, wobei sowohl die Anzahl der Fahrten als auch die durchschnittliche Tripdauer berücksichtigt werden. Im Jahr 2022 erreichte die Nutzung von Fahrrädern ihren Höhepunkt mit einem beeindruckenden Wert von 13.500.000 Fahrten. Von diesem Spitzenjahr an ist jedoch ein deutlicher Rückgang zu verzeichnen, und die Anzahl der Fahrten sinkt drastisch ab. Interessanterweise spiegelt die Dauerkurve der durchschnittlichen Tripdauer einen ähnlichen Verlauf wider, was darauf hindeutet, dass nicht nur die Häufigkeit der Fahrten, sondern auch die durchschnittliche Dauer pro Fahrt abgenommen hat.

Ein besonderes Augenmerk liegt auf dem Jahr 2023, in dem die Zahl der Fahrten auf 5.343.000 zurückgegangen ist. Diese Entwicklung ist besonders relevant, da sie auf eine weiterhin dynamische Veränderung im Fahrradnutzungsverhalten hinweist. Der interaktive Charakter des Graphen ermöglicht es, detaillierte Einblicke in verschiedene Jahre zu erhalten und spezifische Trends genauer zu analysieren. Die Daten bieten somit eine wichtige Grundlage für weitere Untersuchungen und Maßnahmen zur Förderung der Fahrradnutzung in der Region.

Verteilung der Nutzertypen (Subscriber vs. Customer) im Verhältnis zur Gesamtanzahl – Kuchendiagramm

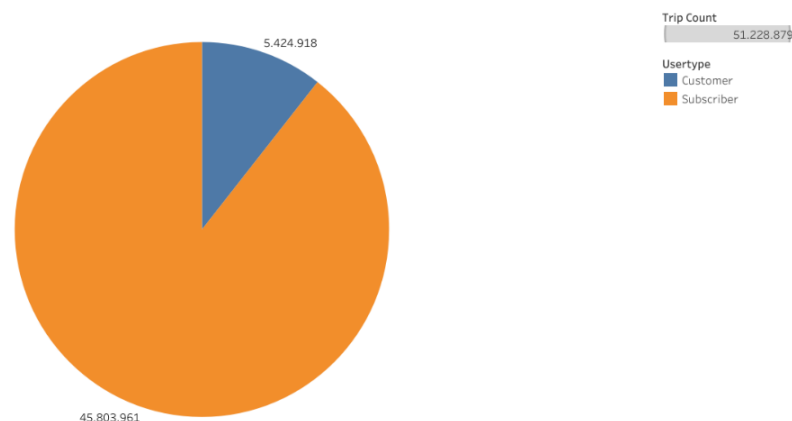


Abbildung 13: Verteilung der Nutzertypen (Subscriber vs. Customer)

(Quelle: Tableau Software – eigene Darstellung)

Das vorliegende Diagramm zeigt die Verteilung der Nutzertypen (Subscriber vs. Customer) im Zeitraum von 2018 bis 2023 in Relation zur Gesamtanzahl der Nutzer.

Das Diagramm ist in Form eines Kuchendiagramms gestaltet, wodurch die Anteile der beiden Nutzertypen anschaulich dargestellt werden.

Im betrachteten Zeitraum ergibt sich, dass von insgesamt 51.228.879 Nutzern 45.803.961 als "Subscriber" klassifiziert sind, was einen Großteil der Gesamtnutzer ausmacht. Dies weist darauf hin, dass eine beträchtliche Anzahl von Nutzern regelmäßige Abonnenten des Fahrradservices ist. Auf der anderen Seite zeigt das Diagramm, dass 5.424.918 Nutzer als "Customer" eingestuft sind. Kunden in diesem Kontext könnten gelegentliche oder erstmalige Nutzer sein, die den Service nicht regelmäßig abonnieren.

Die visuelle Darstellung verdeutlicht somit nicht nur die Gesamtanzahl der Nutzer, sondern auch den prozentualen Anteil der verschiedenen Nutzertypen, was wichtige Erkenntnisse für die Analyse und Planung von Marketing- und Servicestrategien bietet.

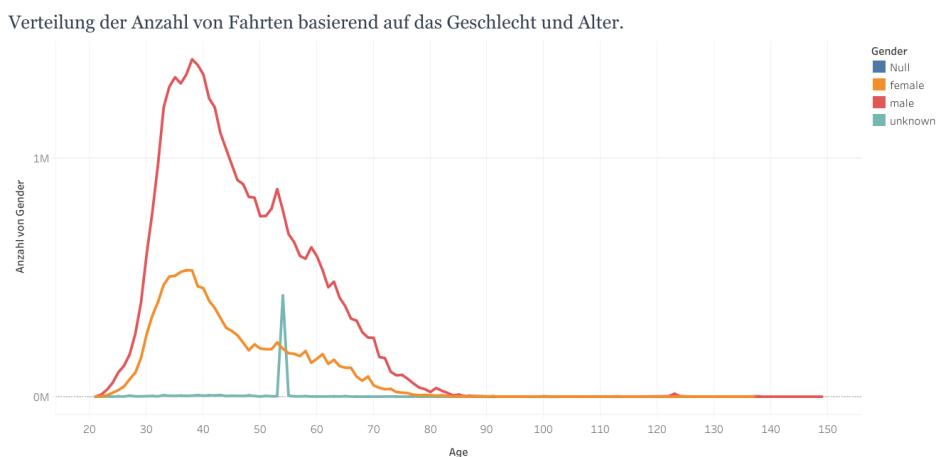


Abbildung 14: Verteilung der Anzahl der Fahrten basierend auf das Geschlecht und Alter

(Quelle: Tableau Software – eigene Darstellung)

Der vorliegende Graph bietet eine detaillierte Darstellung der Fahrradausleihen im Zeitraum von 2018 bis 2023 und legt dabei den Fokus auf die Verteilung nach den Dimensionen Geschlecht und Alter. Hier wurden 58.937.715 Teilnehmer untersucht. Deutlich wird, dass Männer im Vergleich zu Frauen signifikant häufiger auf das Ausleihen von Fahrrädern zurückgreifen. Das Spitzenalter für männliche Nutzer liegt bei 38 Jahren, wobei in diesem Alter beachtliche 1.417.694 Fahrten erfasst wurden. Bei weiblichen Nutzern zeigt sich ein ähnliches Bild, auch hier liegt das Spitzenalter bei 38 Jahren mit 530.665 erfassten Fahrten. Interessanterweise gibt es eine Gruppe von Nutzern, bei denen das Geschlecht unbekannt ist, da sie diese Information nicht angegeben haben. Diese Nutzergruppe ist ebenfalls in der Gesamtanzahl der Fahrten berücksichtigt.

Häufigkeit und Nutzertypen der Startstationen nach Stadtgebiet und Stadtteil

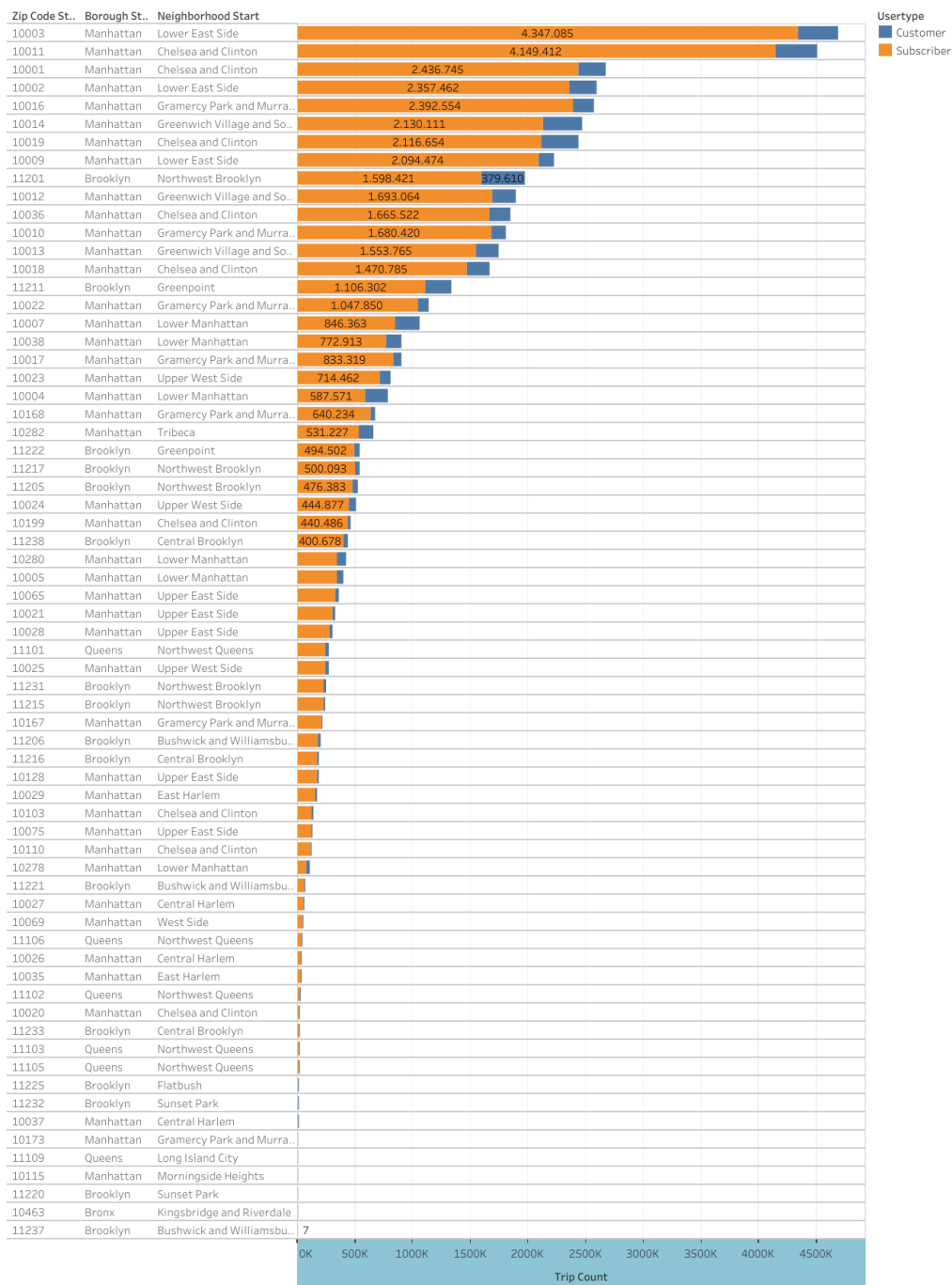


Abbildung 15 : Häufigkeit und Nutzertypen der Startstationen nach Stadtgebiet und Stadtteil

(Quelle: Tableau Software – eigene Darstellung)

Die Grafik 15 liefert eine detaillierte Darstellung der Verteilung von Fahrradnutzungen in verschiedenen Stadtteilen von 2018 bis 2023. In den Stadtgebieten Lower East Side, Chelsea and Clinton, Gramercy Park & Murray Hill, Greenwich Village and SoHo sowie Northwest Brooklyn finden besonders häufig Fahrradfahrten statt.

Die Unterscheidung nach Nutzertypen (Abonnenten und Erstkunden) ermöglicht eine tiefere Analyse der Fahrradnutzungsdynamik in diesen Gebieten. Dieser Ansatz ermöglicht es, festzustellen, ob bestimmte Teile der Stadt von regelmäßigen Abonnenten oder gelegentlichen Erstkunden bevorzugt werden. Die Informationen aus dieser Grafik sind entscheidend für die Stadtplanung und können dazu beitragen, gezielte Maßnahmen zur Förderung der Fahrradmobilität in bestimmten Stadtteilen zu ergreifen.

Häufigkeit und Nutzertypen der Endstationen nach Stadtgebiet und Stadtteil

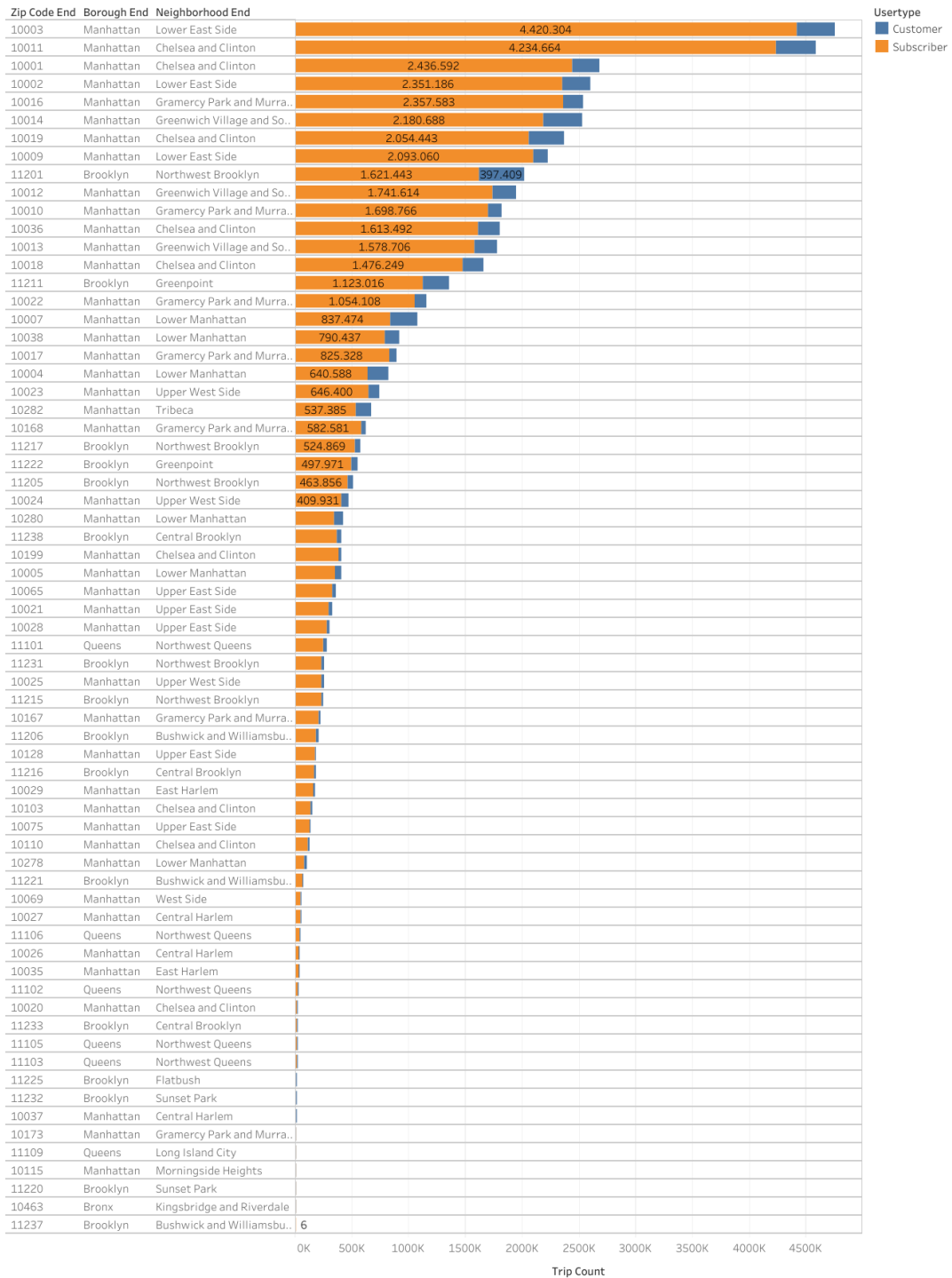


Abbildung 16 : Häufigkeit und Nutzertypen der Endstationen nach Stadtgebiet und Stadtteil

(Quelle : Tableau Software – eigene Darstellung)

Die Grafik bietet einen detaillierten Einblick in die Fahrradnutzung von 2018 bis 2023, indem sie die Endstationen in verschiedenen Stadtteilen analysiert. Die Daten zeigen, dass bestimmte Stadtgebiete besonders häufig als Ziel für Fahrradfahrten ausgewählt werden. Dazu zählen Lower East Side, Chelsea and Clinton, Gramercy Park & Murray Hill, Greenwich Village and SoHo sowie Northwest Brooklyn.

Die Betrachtung des Nutzertyps verleiht der Grafik zusätzliche Aussagekraft. Sie differenziert zwischen Abonnenten und Erstkunden, was wertvolle Einblicke in die Präferenzen und Verhaltensmuster verschiedener Nutzergruppen ermöglicht. Diese Erkenntnisse sind nicht nur für städtische Planer relevant, um gezielte Verbesserungen der Fahrradinfrastruktur vorzunehmen, sondern auch für Unternehmen im Mobilitätssektor, um ihre Dienstleistungen besser auf die Bedürfnisse der Nutzer abzustimmen.

Die Grafik verdeutlicht somit die dynamische Entwicklung der Fahrradnutzung in verschiedenen Stadtteilen und zeigt, wie der Nutzertyp die Auswahl der Endstationen beeinflusst. Dies ermöglicht fundierte Entscheidungen zur Förderung nachhaltiger Mobilität und zur Optimierung von Fahrradinfrastrukturen.

Einfluss der Niederschlagswahrscheinlichkeit auf die Anzahl der Fahrten

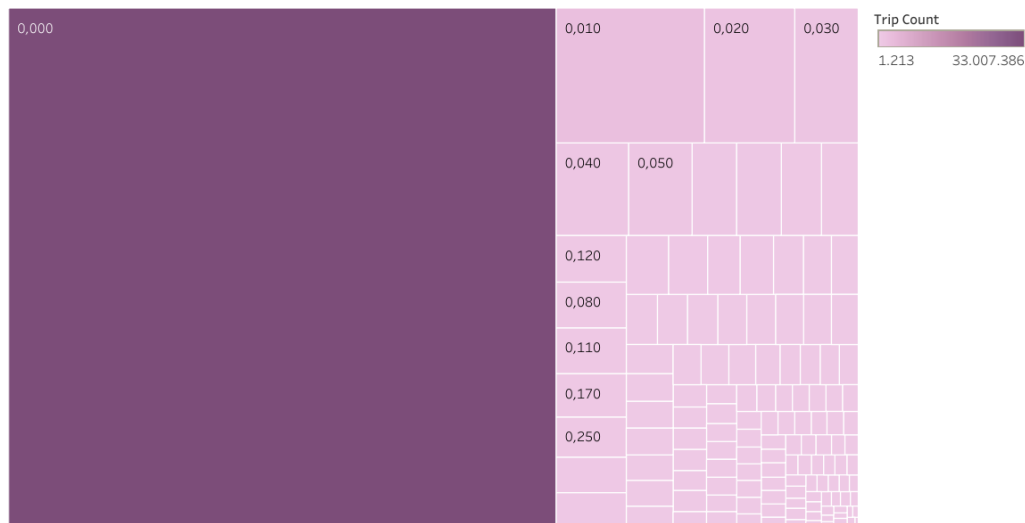


Abbildung 17 : Einfluss der Niederschlagswahrscheinlichkeit auf die Anzahl der Fahrten

(Quelle : Tableau Software – eigene Darstellung)

Die Grafik für den Zeitraum von 2018 bis 2023 bietet einen aufschlussreichen Überblick über das Verhältnis zwischen der Wahrscheinlichkeit von Niederschlägen und der Anzahl der Fahrradfahrten. Im Jahr 2022 verzeichnete die Grafik die höchste Anzahl von Fahrten, wobei eine auffällige Beziehung zur Niederschlagswahrscheinlichkeit zu erkennen ist.

Das Kastendiagramm verdeutlicht, dass bei einer Niederschlagswahrscheinlichkeit von 0 die Fahrradnutzung besonders hoch ist. Mit steigender Niederschlagswahrscheinlichkeit nimmt die Anzahl der Fahrten ab. Diese negative Korrelation zwischen Niederschlagswahrscheinlichkeit und Fahrradnutzung liefert wertvolle Einblicke.

Die Entscheidung für ein Heatmap-Format ermöglicht es, die dargestellten Zahlen effektiv zu visualisieren. Die Farbintensität im Heatmap zeigt die Intensität der Beziehung zwischen Niederschlagswahrscheinlichkeit und Fahrtenanzahl an. Dabei wurde bewusst auf eine anschauliche und leicht interpretierbare Darstellung gesetzt, um komplexe Zusammenhänge verständlich zu machen.

Korrelation zwischen durchschnittlicher Tagestemperatur und Anzahl der Fahrten

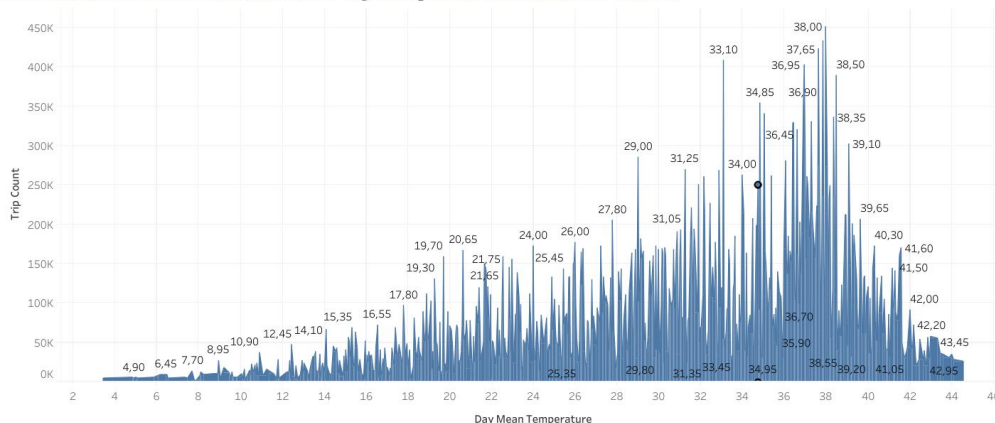


Abbildung 18: Korrelation zwischen durchschnittlicher Tagestemperatur und Anzahl der Fahrten

(Quelle: Tableau Software – eigene Darstellung)

Die Visualisierung der Korrelation zwischen der durchschnittlichen Tagestemperatur und der Anzahl der Fahrten im gesamten Zeitraum bietet interessante Einblicke in das Verhalten der Fahrradnutzung. Auffällig ist, dass die höchste Anzahl von Fahrten bei bestimmten Temperaturbereichen liegt, insbesondere bei 36, 37, 38 und 33 Grad Celsius.

Die Grafik verdeutlicht, dass es eine deutliche positive Korrelation zwischen der Temperatur und der Anzahl der Fahrradfahrten gibt. An Tagen mit den genannten Temperaturen nutzen die Menschen offenbar vermehrt Fahrräder. Dies könnte auf die angenehmen Wetterbedingungen zurückzuführen sein, die das Radfahren attraktiver machen.

Die Wahl von bestimmten Temperaturbereichen als Peaks in der Fahrradnutzung könnte für städtische Planer und Mobilitätsdienstleister von Bedeutung sein. Es könnte beispielsweise Überlegungen zur Verbesserung der Fahrradinfrastruktur oder zur Förderung von Fahrradnutzung an Tagen mit ähnlichen Temperaturen anregen.

Die klare Darstellung der Korrelation in dieser Grafik ermöglicht es, Muster und Trends leicht zu erkennen und fundierte Schlussfolgerungen über den Einfluss der Temperatur auf die Fahrradnutzung zu ziehen.

Beziehung zwischen durchschnittlicher Windgeschwindigkeit (m/s) und Anzahl der Fahrten

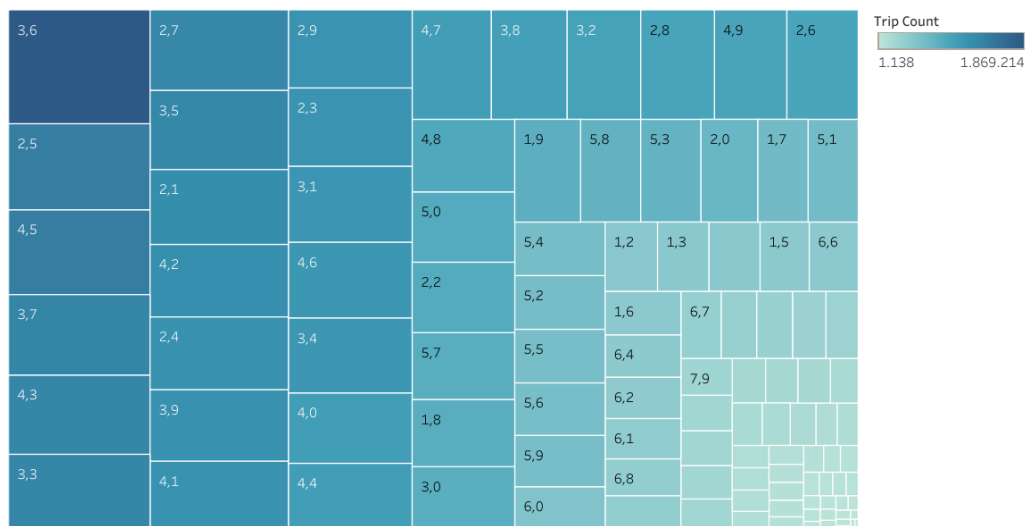


Abbildung 19: Beziehung zwischen durchschnittlicher Windgeschwindigkeit (m/s) und Anzahl der Fahrten

(Quelle: Tableau Software – eigene Darstellung)

Die Visualisierung der Korrelation zwischen durchschnittlicher Windgeschwindigkeit und der Anzahl der Fahrten über den gesamten Zeitraum bietet einen Einblick in das Fahrradnutzungsverhalten in Abhängigkeit von der Windgeschwindigkeit. Auffällig ist, dass die höchste Anzahl von Fahrten bei einer Windgeschwindigkeit von 3,6 m/s registriert wurde. Diese Geschwindigkeit entspricht dem Punkt, an dem Blätter und dünnere Zweige sich bewegen und Wimpel sich strecken.

Die Verwendung einer Heatmap zur Übersicht stellt die Daten anschaulich dar: Je intensiver die Farbe in den Zellen ist, desto mehr Fahrten wurden in Bezug auf die spezifische Windgeschwindigkeit verzeichnet. Diese visuelle Darstellung ermöglicht eine schnelle Identifizierung von Bereichen mit besonders hoher oder niedriger Fahrradaktivität in Abhängigkeit von der Windgeschwindigkeit.

Eine gezielte Berücksichtigung der Windverhältnisse könnte in Maßnahmen zur Förderung der Fahrradmobilität einfließen, um beispielsweise besonders windige Strecken attraktiver zu gestalten.

4.2.3.5.2 Datenvisualisierung, /-Analyse für das Sommer-Saison 2022

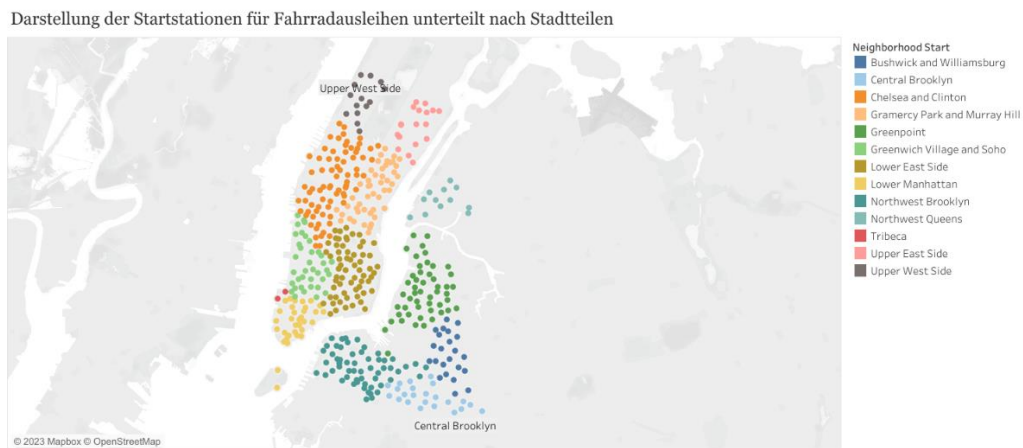


Abbildung 20: Darstellung der Startstationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken

(Quelle: Tableau Software – eigene Darstellung)

Die geografische Visualisierung der Fahrradnutzung in New York City innerhalb eines Stadtgebiets bietet einen effizienten Ansatz zur Entdeckung geografischer Muster und Hotspots im Zusammenhang mit der Fahrradnutzung. Die bunten Punkte in der Darstellung repräsentieren unterschiedliche Stadtgebiete und haben eine standardisierte Größe, da sie Fahrradstationen symbolisieren, die als Ausgangspunkte für Fahrradtouren dienen.

Jede dieser Stationen gilt als Ausgangspunkt für Fahrradtouren, wobei einige Stationen noch keine zeitmäßigen Aufzeichnungen aufweisen.

Dies kann wertvolle Erkenntnisse für die städtische Planung und die Verbesserung des Fahrradverkehrs in der Stadt New York liefern, indem sie aufzeigt, welche Stadtteile besonders beliebt für Fahrradfahrten sind und wo möglicherweise zusätzliche Ressourcen für Fahrradstationen oder Infrastruktur benötigt werden.

In der nächsten Abbildung wird die Skalierung der Kreise anhand der Dauer der Fahrt an jeder Station veranschaulicht. Dies ermöglicht eine visuelle Darstellung der Variationen in der Nutzungsdauer von Fahrrädern an verschiedenen Standorten. Größere Kreise repräsentieren längere durchschnittliche Fahrradtouren an den jeweiligen Stationen.

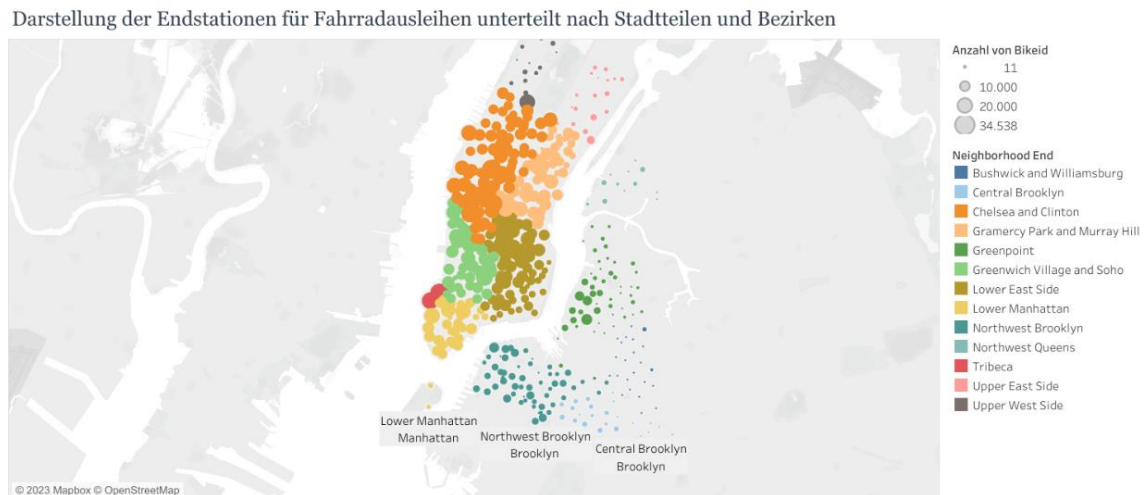


Abbildung 21: Darstellung der Abgabestationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken

(Quelle: Tableau Software – eigene Darstellung)

Die Karte ist eine effektive Möglichkeit, um geografische Muster und Hotspots in Bezug auf die Nutzung von Fahrrädern zu identifizieren.

Merkmale der Grafik:

Farbliche Kennzeichnung: Die Stadtteile und Bezirke sind durch unterschiedliche Farben gekennzeichnet, um eine klare Unterscheidung zu ermöglichen. Jede Farbe repräsentiert einen bestimmten Bereich.

Kreise für Ausleihen: An den Startstationen sind Kreise platziert, deren Größe proportional zur Anzahl der ausgeliehenen Fahrräder ist. Größere Kreise zeigen stark frequentierte Stationen an, während kleinere Kreise auf weniger genutzte Stationen hinweisen.

Interaktivität: Die Karte kann interaktiv sein, was bedeutet, dass Benutzer mit ihr interagieren können. Zum Beispiel können sie auf einen bestimmten Kreis klicken, um detaillierte Informationen zur jeweiligen Startstation anzuzeigen.

Die Karte bietet einen umfassenden Überblick über die räumliche Verteilung der Fahrradnutzung. Entscheidungsträger können schnell erkennen, welche Stadtteile die Hauptnutzungsbereiche sind und wie sich die Ausleihe auf die verschiedenen Bezirke verteilt.

4.2.4 Microsoft Azure + Power BI

Im vorherigen Schritt, insbesondere während der Umsetzung in der GCP-Technologie, wurden die Daten bereits gründlich bereinigt und strukturiert. In dieser aktuellen Implementierung in Microsoft Azure werden jedoch keine Transformationsprozesse durchgeführt. Stattdessen erfolgen lediglich die Prozesse des Ladens (Load) und Extrahierens (Extract).

Diese Entscheidung wurde getroffen, um die Effizienz des Datenflusses zu optimieren und die Gesamtumsetzungszeit zu verkürzen. Da die Daten bereits in einer geeigneten Form vorliegen, ist es nicht erforderlich, weitere Transformationsprozesse durchzuführen, die die Daten in ihrem Kern verändern würden. Der Fokus liegt daher darauf, die bereinigten und strukturierten Daten aus der GCP in Microsoft Azure zu laden und zu extrahieren, um nahtlose Interoperabilität zwischen den Plattformen zu gewährleisten.

4.2.4.1 Azure Analysis Services

Der Begriff Cloud Computing gewinnt zunehmend an Bedeutung und Interesse bei vielen Unternehmen. Anbieter wie Microsoft, Google oder Amazon machen Cloud Computing wettbewerbsfähiger und attraktiver sowohl für private Nutzer als auch für Unternehmen. „Diese Cloud Service Provider verteilen riesige Rechenzentren über den gesamten Globus, um ihren Kunden einen lokalen Service mit niedrigen Latenzzeiten bieten zu können“.⁵⁷ Cloud Computing hat sich in den letzten Jahren zu einem Wirtschaftszweig entwickelt.⁵⁸

Cloud Computing bietet im Prinzip Flexibilität bei der Bereitstellung, Nutzung und Abbestellung von IT-basierten Diensten. Dabei werden verschiedene Abrechnungsmodelle angeboten, wie: Pay-as-you-go oder On-Demand. Hierbei bezahlt der Kunde nur für diejenigen Ressourcen, die er benutzt.⁵⁹

Microsoft Azure Analysis Services gilt als leistungsstarke und renommierte Cloud-Lösung für die Daten-Analyse, Datenmodellierung und -verwaltung. Diese wird unter Saas (Software as a Service) kategorisiert.⁶⁰

Die vorliegende Arbeit widmet sich sowohl einer Untersuchung der Funktionalitäten von Azure Analysis Services, einem integralen Bestandteil der Microsoft Azure Cloud-Plattform, als auch der Herstellung einer Verbindung mit Power BI. In Anbetracht der wachsenden Bedeutung von Datenanalysen in heutigen Unternehmensumgebungen ist das Verständnis und die effektive Nutzung von leistungsstarken Analysewerkzeugen entscheidend.

⁵⁷ (vgl. Kohne 2018, S. 2)

⁵⁸ (vgl. ebd.: S. 1)

⁵⁹ (vgl. ebd.: S. 1)

⁶⁰ (Was ist Azure Analysis Services? o. D.)

Der vorliegende Screenshot zeigt den Prozess der Einrichtung des "Cyclistic"-Verzeichnisses in Microsoft Analysis Services. Dieser Schritt ist entscheidend, um eine geordnete Struktur für die Datenanalyse und -verarbeitung zu schaffen. Das Verzeichnis enthält relevante Datensätze und Metadaten, die für die Analyse und Generierung von Einblicken von großer Bedeutung sind.

Zusätzlich zur Einrichtung des Verzeichnisses wurde auch ein Server in der Cloud konfiguriert, um die Leistung und Skalierbarkeit des Systems sicherzustellen. Dieser Schritt ermöglicht den nahtlosen Zugriff auf die erforderlichen Ressourcen für die Datenverarbeitung und -analyse.

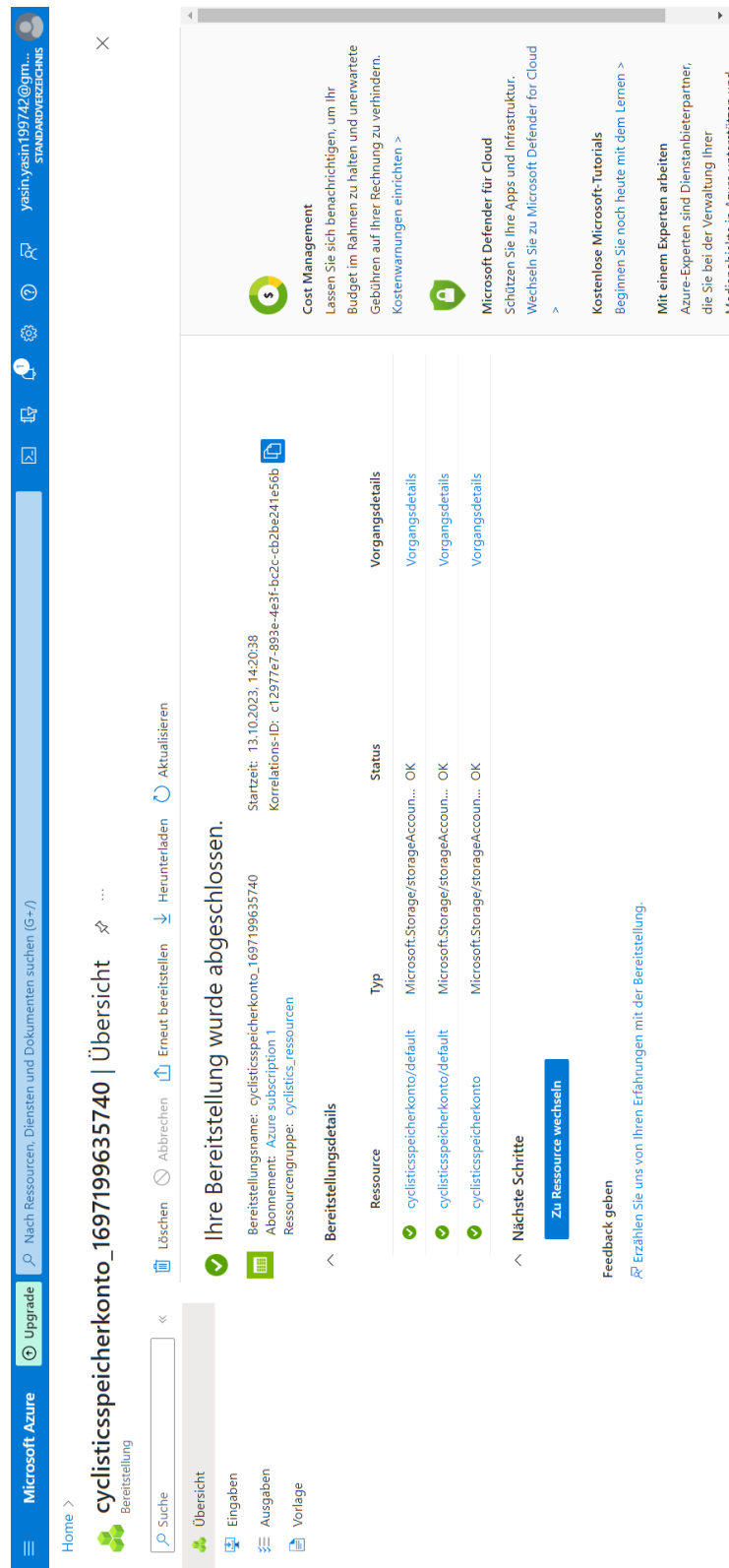


Abbildung 23: Einrichtung des „Cyclistic“-Verzeichnisses in Microsoft Analysis Services

Der vorliegende Screenshot veranschaulicht einen Schritt im Rahmen der Arbeit mit Microsoft Analysis Services, nämlich die Erstellung eines Speicherkontos in einem Verzeichnis mit dem Namen "Cyclistic". Dieser Schritt ist von großer Bedeutung, da er die Grundlage für die Datenverwaltung und -integration bildet.

Auf dem Screenshot ist zu erkennen, wie das Speicherkonto mit dem Namen "cyclisticspeicherkonto" innerhalb des Verzeichnisses "Cyclistic" in Microsoft Analysis Services erstellt wurde. Das Speicherkonto fungiert als Schnittstelle für den Zugriff auf externe Datenquellen, in diesem Fall mit einem spezifischen Fokus auf die Datenspeicherung in der Cloud.

Die Einrichtung des Speicherkontos wurde unter Berücksichtigung der Projekterfordernisse vorgenommen und umfasst die Festlegung von Sicherheitseinstellungen, Übertragungsprotokollen und Speicherkapazitäten. Dies gewährleistet nicht nur den sicheren Zugriff auf die Daten, sondern auch die effiziente Verwaltung und Übertragung von Informationen.

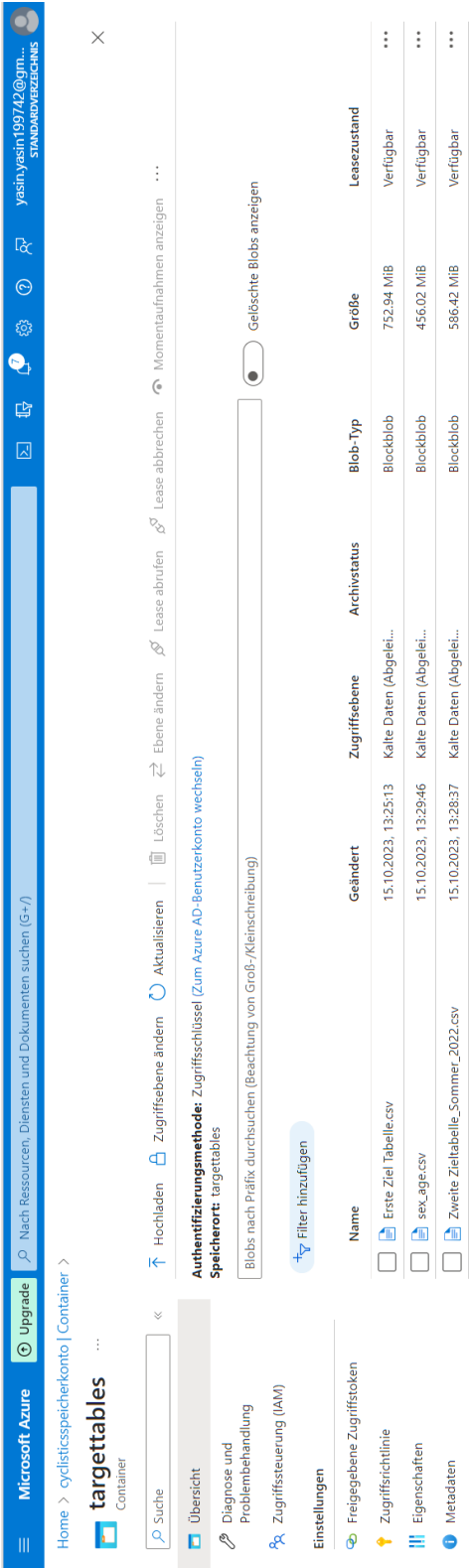


Abbildung 24: Einrichtung des „Cyclistic“-Verzeichnisses in Microsoft Analysis Services

(Quelle: Microsoft Azure – eigene Darstellung)

Der vorliegende Screenshot visualisiert einen kritischen Schritt in der Verwendung von Microsoft Analysis Services, nämlich das Hochladen der drei bereinigten Ziel-Tabellen. Diese Tabellen wurden im vorherigen Datenbereinigungsprozess auf der Google Cloud Plattform erstellt und sind von entscheidender Bedeutung für die folgende Datenanalyse.

Der Screenshot zeigt, wie diese Ziel-Tabellen erfolgreich in Microsoft Analysis Services hochgeladen wurden. Das Speicherkonto, mit dem Namen "cyclisticspeicherkonto," das zuvor im Rahmen des Projekts konfiguriert wurde, dient als Schnittstelle für den Datentransfer zwischen der Google Cloud Plattform und Microsoft Analysis Services.

Der Hochladevorgang selbst wurde sorgfältig geplant und umgesetzt, um sicherzustellen, dass die Datenintegrität gewahrt bleibt und die Daten effizient in Microsoft Analysis Services verwendet werden können.

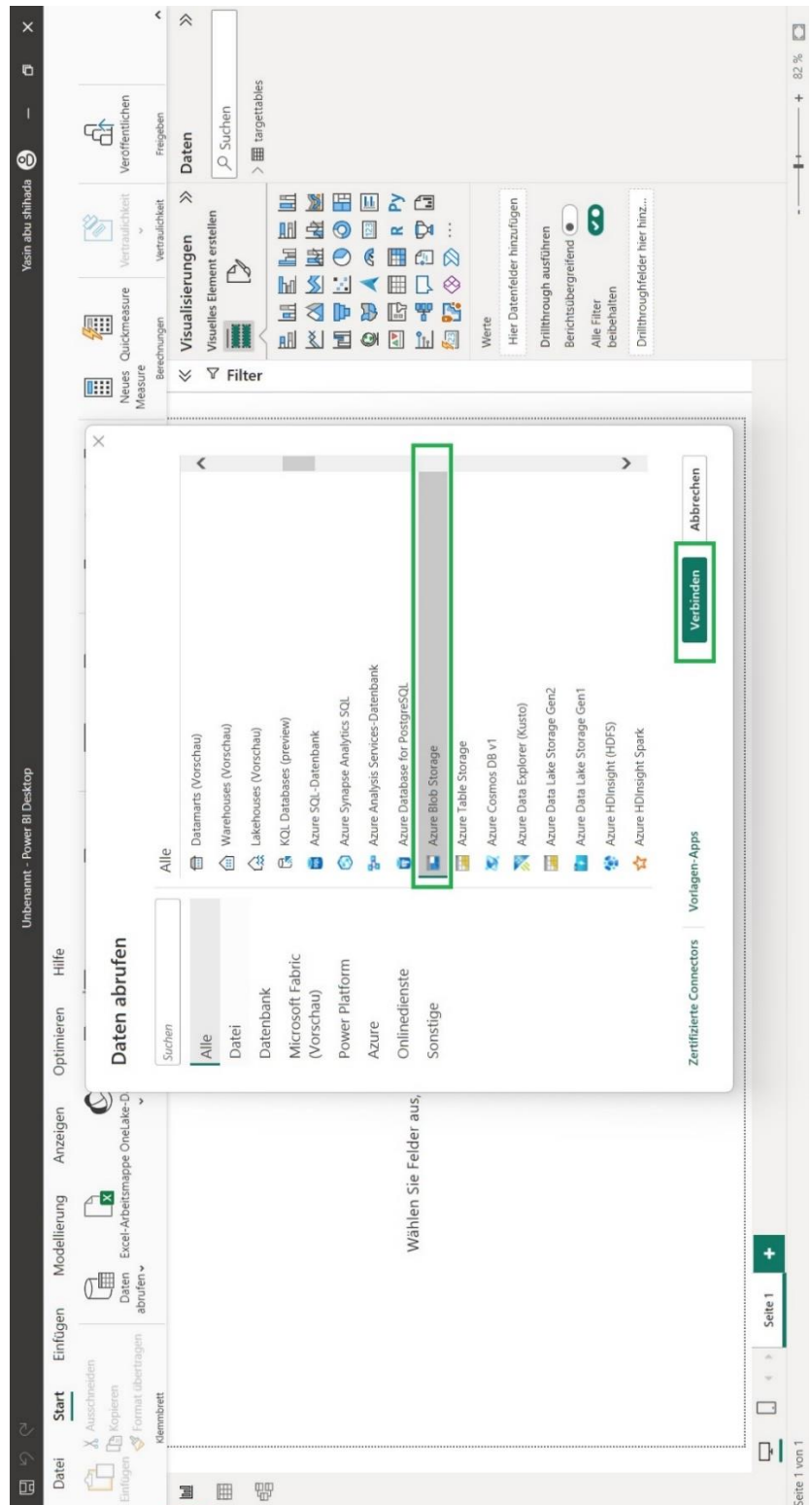


Abbildung 25: Integration von Daten aus Azure Blob Storage in Power BI

(Quelle: Power BI – eigene Darstellung)

Die Integration von Daten aus Azure Blob Storage in Power BI ist ein weiterer Schritt im Datenmanagement und der Analyse. Azure Blob Storage bietet eine zuverlässige und skalierbare Möglichkeit zur Speicherung von Daten in der Cloud. Power BI hingegen ist eine leistungsstarke Business-Intelligence-Plattform, die es ermöglicht, Daten zu visualisieren und fundierte Geschäftsentscheidungen zu treffen.

Die Integration dieser beiden Plattformen eröffnet vielfältige Möglichkeiten. Zunächst erfolgt die Verbindung mit Azure Blob Storage, um auf die dort gespeicherten Daten zuzugreifen. Dieser Schritt ermöglicht den Datenaustausch zwischen der Cloud und der Business-Intelligence-Plattform ermöglicht.

Sobald die Verbindung hergestellt ist, können die Daten aus Azure Blob Storage in Power BI importiert und modelliert werden. Dies umfasst die Auswahl der relevanten Datenquellen und die Anwendung von Transformationen, um die Daten für die Analyse vorzubereiten.

Die Daten können dann in Power BI verwendet werden, um ansprechende Berichte und Visualisierungen zu erstellen. Mit den integrierten Daten können Benutzer Muster und Trends erkennen, die in den Informationen verborgen sind.

Die Integration von Azure Blob Storage und Power BI bietet Organisationen die Möglichkeit, Daten effizient zu nutzen und Geschäftseinblicke zu gewinnen. Dieser Prozess ermöglicht es, auf eine breite Palette von Cloud-basierten Datenquellen zuzugreifen und sie in aussagekräftige Berichte und Visualisierungen umzuwandeln.

4.2.4.2 Datenvisualisierung & Gewinnung von Erkenntnissen

4.2.4.2.1 Datenvisualisierung-/Analyse für den Zeitraum (2018-2022)



Abbildung 26: Jahresvergleich: Fahrradnutzung und Tripdauer

(Quelle: Power BI Software – eigene Darstellung)

Die Erstellung eines Dashboards mit Power BI ermöglicht die übersichtliche und effektive Darstellung von Daten. In diesem speziellen Dashboard werden verschiedene wichtige Metriken für die Jahre 2018 bis 2023 zusammengefasst.

Das Dashboard zeigt zuerst die Gesamtsumme der Fahrten nach Jahr im Zeitraum (2018-2023), was einen Überblick über die jährliche Entwicklung bietet. Somit kann man die Trends und Veränderungen im Laufe der Jahre identifizieren.

Zusätzlich zur jährlichen Aufschlüsselung wird die Gesamtsumme der Fahrtminuten nach Jahr von 2018 bis 2023 dargestellt. Dies ermöglicht es, die gesamte Dauer der Fahrten im Zeitverlauf zu analysieren und gegebenenfalls Zusammenhänge zu identifizieren.

Das Dashboard geht jedoch über die jährliche Analyse hinaus und zeigt auch die Summe der Fahrten nach Monat. Die monatliche Aufschlüsselung bietet einen tieferen Einblick in die Variation der Fahrten über das Jahr hinweg und hilft dabei, saisonale Muster oder Veränderungen in bestimmten Monaten zu erkennen.

Zusätzlich zu den oben genannten Metriken zeigt das Dashboard auch die Gesamtsumme der Fahrten und Fahrtminuten.

Summe von gesamten Fahrten nach usertype

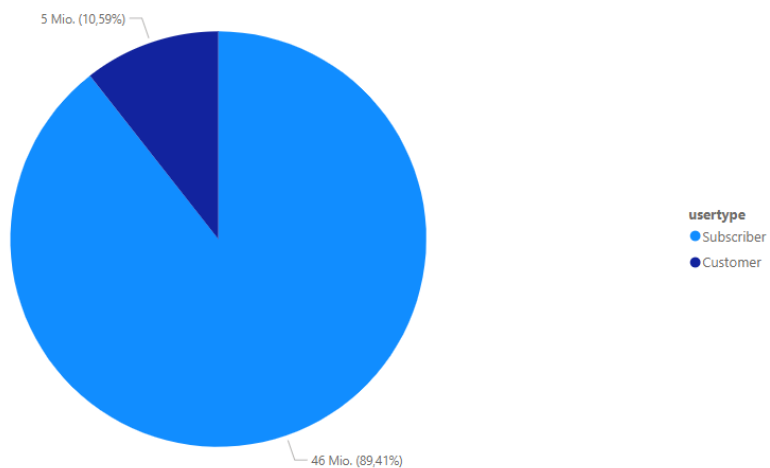


Abbildung 27: Verteilung der Nutzertypen (Subscriber vs. Customer)

(Quelle: Power BI Software – eigene Darstellung)

Das vorliegende Diagramm wurde mithilfe von Power BI erstellt und zeigt die Verteilung der Nutzertypen (Subscriber vs. Customer) im Zeitraum von 2018 bis 2023 in Relation zur Gesamtanzahl der Nutzer. Das Diagramm wurde als Kuchendiagramm gestaltet, um die Anteile der beiden Nutzertypen auf anschauliche Weise darzustellen.

Während des analysierten Zeitraums wurden insgesamt 51.228.879 Nutzer erfasst. Von diesen Nutzern wurden 45.803.961 als "Subscriber" klassifiziert, was den Großteil der Gesamtnutzer ausmacht. Dies deutet darauf hin, dass eine erhebliche Anzahl von Nutzern den Fahrradservice regelmäßig abonniert. Andererseits wurden 5.424.918 Nutzer als "Customer" eingestuft. In diesem Zusammenhang könnten Kunden gelegentliche oder erstmalige Nutzer sein, die den Service nicht regelmäßig abonnieren.

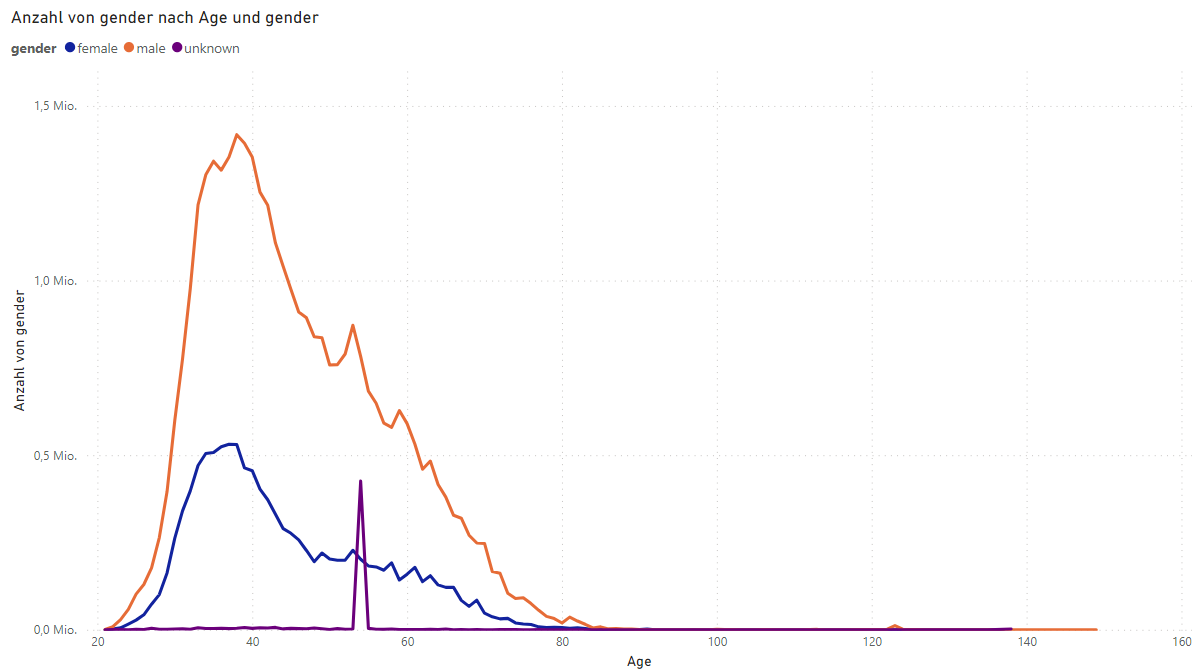


Abbildung 28: Verteilung der Anzahl der Fahrten basierend auf das Geschlecht und Alter

(Quelle: Power BI Software – eigene Darstellung)

Das vorliegende Diagramm präsentiert eine umfassende Abbildung der Fahrradausleihen im Zeitraum von 2018 bis 2023, wobei der Fokus auf der Segmentierung nach Geschlecht und Altersgruppen liegt. In dieser Analyse wurden 58.937.715 Teilnehmer berücksichtigt. Deutlich wird, dass männliche Nutzer im Vergleich zu ihren weiblichen Nutzern signifikant häufiger das Fahrradverleihsystem in Anspruch nehmen. Das höchste Nutzeralter bei männlichen Nutzern liegt bei 38 Jahren, und in dieser Altersgruppe wurden beachtliche 1.417.694 Fahrten registriert. Eine ähnliche Beobachtung zeigt sich bei weiblichen Nutzern, bei denen ebenfalls das höchste Nutzeralter bei 38 Jahren liegt, und 530.665 Fahrten in dieser Altersgruppe erfasst wurden. Interessanterweise gibt es eine Gruppe von Nutzern, bei denen das Geschlecht nicht bekannt ist, da diese Informationen nicht angegeben wurden. Diese Nutzergruppe ist ebenfalls in die Gesamtzahl der erfassten Fahrten einbezogen.

Summe von gesamten Fahrten nach PLZ und Anfangsstationen

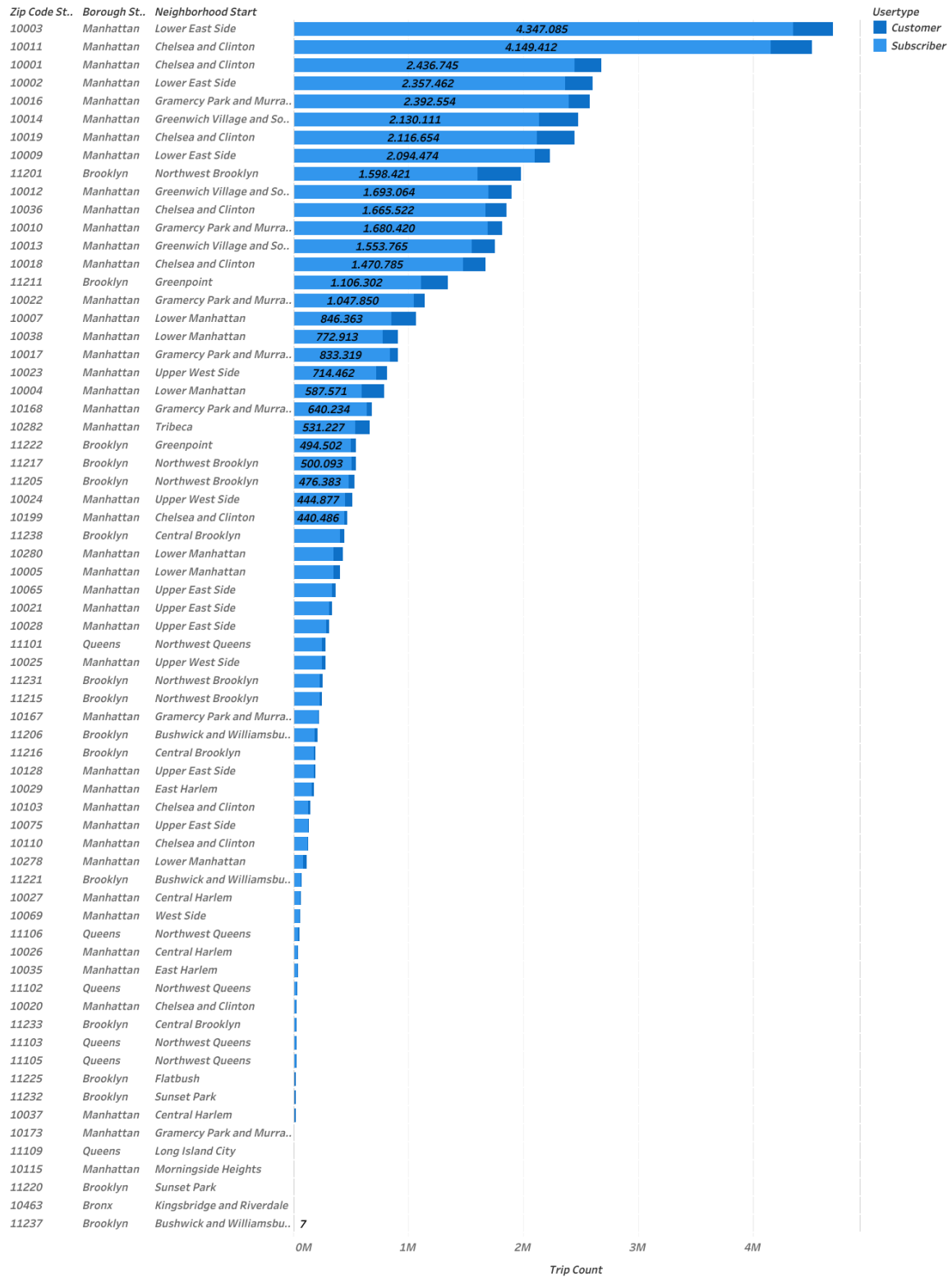


Abbildung 29: Häufigkeit und Nutzertypen der Startstationen nach Stadtgebiet und Stadtteil

(Quelle: Power BI Software – eigene Darstellung)

Die visuelle Darstellung bietet eine umfassende Abbildung der Verteilung der Fahrradnutzung in verschiedenen Stadtvierteln im Zeitraum von 2018 bis 2023. Es zeigt, dass in den Stadtteilen Lower East Side, Chelsea and Clinton, Gramercy Park & Murray Hill, Greenwich Village and SoHo sowie Northwest Brooklyn besonders häufig Fahrradfahrten unternommen werden.

Die Unterscheidung zwischen Nutzertypen, nämlich Abonnenten und erstmaligen Kunden, ermöglicht eine tiefgehende Analyse der Dynamik der Fahrradnutzung in diesen Gebieten. Dieser Ansatz erlaubt es, festzustellen, ob bestimmte Teile der Stadt von regelmäßigen Abonnenten oder sporadischen Erstkunden bevorzugt werden. Die Erkenntnisse aus dieser Darstellung sind von entscheidender Bedeutung für die städtische Planung und können dazu beitragen, gezielte Maßnahmen zur Förderung der Fahrradmobilität in spezifischen Stadtvierteln zu entwickeln.

Summe von gesamten Fahrten nach PLZ und Endstationen

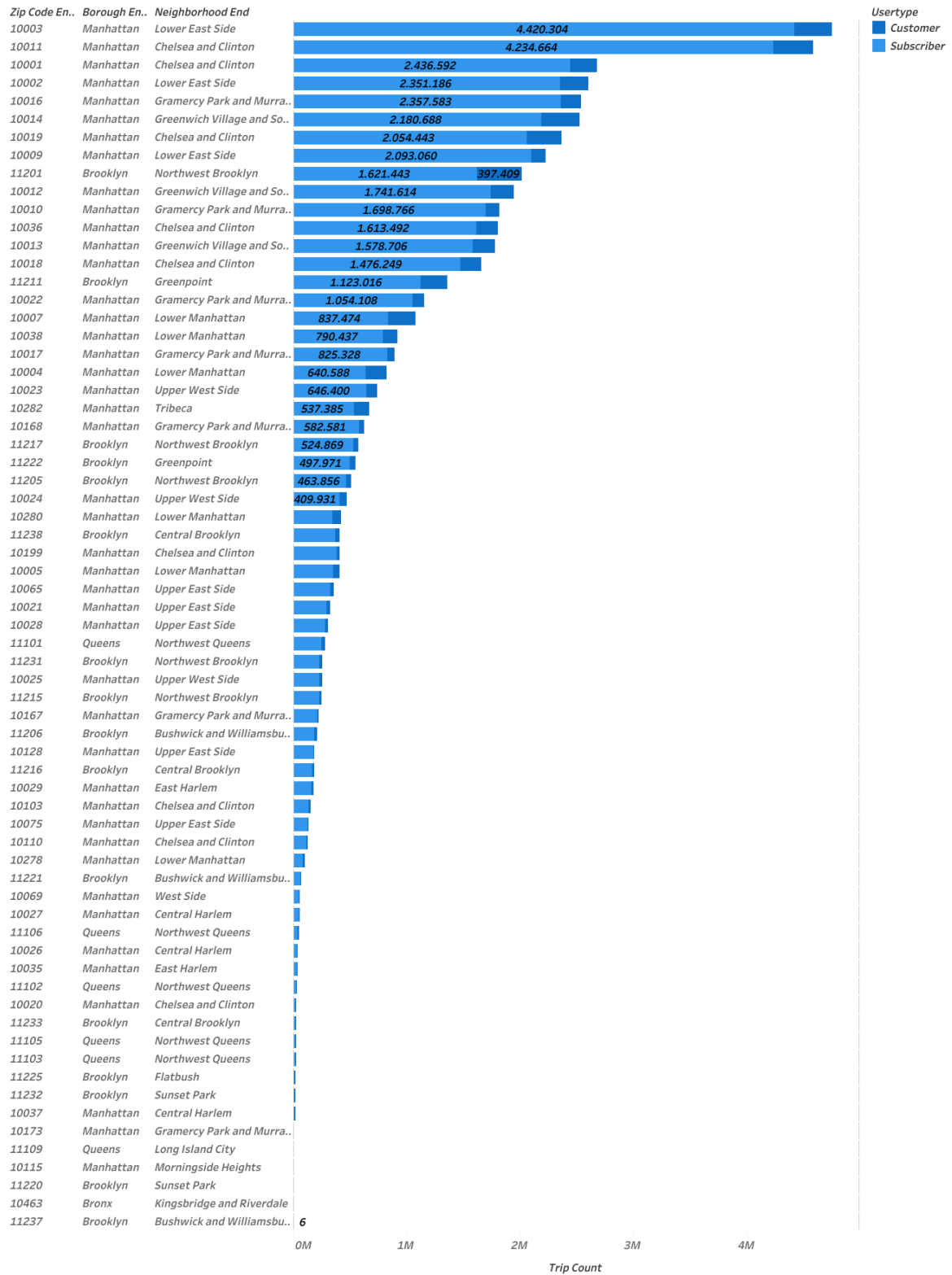


Abbildung 30: Häufigkeit und Nutzertypen der Endstationen nach Stadtgebiet und Stadtteil

(Quelle: Power BI Software – eigene Darstellung)

Die visuelle Darstellung gewährt einen tiefen Einblick in die Fahrradnutzung von 2018 bis 2023, indem sie die Endstationen in verschiedenen Stadtteilen analysiert. Die Daten verdeutlichen, dass bestimmte Stadtviertel besonders häufig als Ziel für Fahrradfahrten ausgewählt werden, darunter Lower East Side, Chelsea and Clinton, Gramercy Park & Murray Hill, Greenwich Village and SoHo sowie Northwest Brooklyn.

Die Berücksichtigung des Nutzertyps verleiht der Grafik zusätzliche Bedeutung, da sie zwischen Abonnenten und erstmaligen Kunden unterscheidet. Dies ermöglicht wertvolle Einblicke in die Präferenzen und Verhaltensmuster unterschiedlicher Nutzergruppen. Diese Erkenntnisse sind nicht nur für städtische Planer von Interesse, um gezielte Verbesserungen der Fahrradinfrastruktur vorzunehmen, sondern auch für Unternehmen im Mobilitätssektor, um ihre Dienstleistungen besser an die Bedürfnisse der Nutzer anzupassen.

Insgesamt verdeutlicht die Grafik die dynamische Entwicklung der Fahrradnutzung in verschiedenen Stadtteilen und zeigt, wie der Nutzertyp die Auswahl der Endstationen beeinflusst.

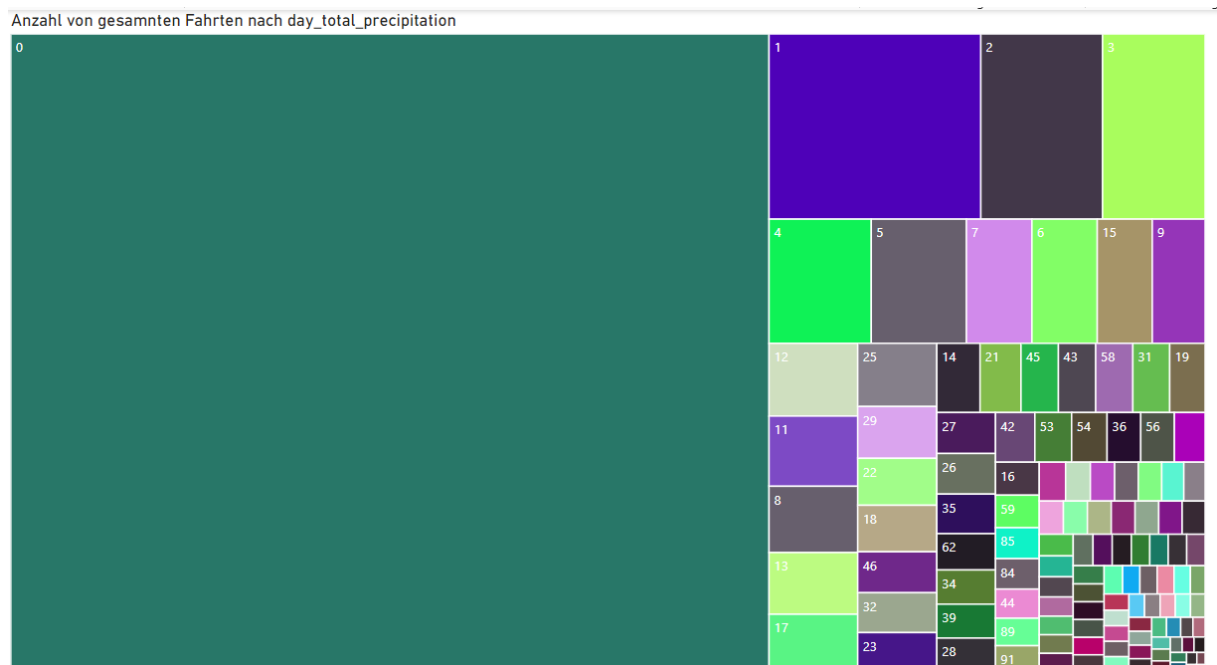


Abbildung 31: Einfluss der Niederschlagswahrscheinlichkeit auf die Anzahl der Fahrten

(Quelle: Power BI Software – eigene Darstellung)

Die Grafik, die den Zeitraum von 2018 bis 2023 abdeckt, bietet einen Überblick über das Verhältnis zwischen der Wahrscheinlichkeit von Niederschlägen und der Anzahl der Fahrradfahrten. Im Jahr 2022 wurden die meisten Fahrten verzeichnet, wobei eine deutliche Verbindung zur Wahrscheinlichkeit von Niederschlägen zu erkennen ist.

Das Diagramm veranschaulicht, dass bei einer Niederschlagswahrscheinlichkeit von 0 die Fahrradnutzung besonders hoch ist. Mit steigender Niederschlagswahrscheinlichkeit nimmt die Anzahl der Fahrten ab. Diese inverse Korrelation zwischen der Wahrscheinlichkeit von Niederschlägen und der Fahrradnutzung liefert wertvolle Erkenntnisse für den Anbieter von Mobilitätsdiensten.

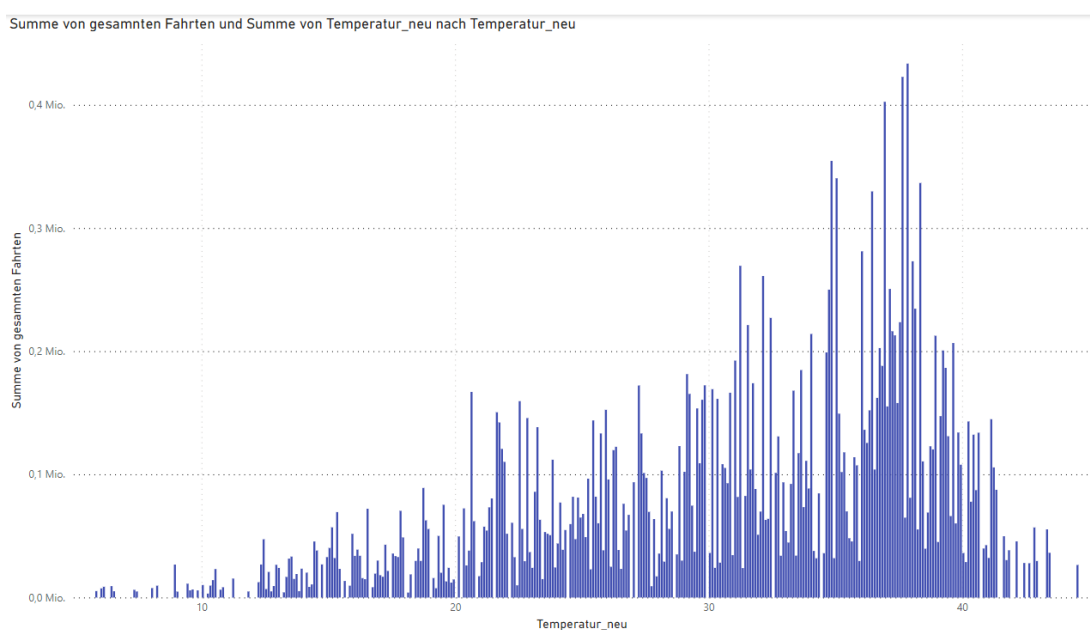


Abbildung 32 : Korrelation zwischen durchschnittlicher Tagestemperatur und Anzahl der Fahrten

(Quelle : Power BI Software – eigene Darstellung)

Die Visualisierung der Wechselbeziehung zwischen der durchschnittlichen Tagestemperatur und der Anzahl der Fahrten im gesamten Zeitraum gewährt faszinierende Einblicke in das Verhalten der Fahrradnutzung.

Besonders auffällig ist, dass die höchste Anzahl an Fahrten bei spezifischen Temperaturbereichen, insbesondere bei 36, 37, 38 und 33 Grad Celsius, zu verzeichnen ist. Die Grafik verdeutlicht eine klare und positive Korrelation zwischen der Temperatur und der Anzahl der Fahrradfahrten.

Es wird ersichtlich, dass an Tagen mit den zuvor genannten Temperaturen die Nutzung von Fahrrädern signifikant zunimmt. Dies könnte auf die angenehmen Wetterbedingungen zurückzuführen sein, die das Fahrradfahren besonders attraktiv gestalten.

Die Festlegung dieser bestimmten Temperaturbereiche als Spitzenwerte in der Fahrradnutzung hat möglicherweise Relevanz für städtische Planer und Anbieter von Mobilitätsdiensten. Es könnte Anreize bieten, die Fahrradinfrastruktur zu optimieren oder die Nutzung von Fahrrädern an Tagen mit ähnlichen Temperaturen zu fördern.

Die klare und prägnante Darstellung dieser Korrelation in der Grafik erleichtert die Identifizierung von Mustern und Trends sowie die Formulierung fundierter Schlussfolgerungen hinsichtlich des Einflusses der Temperatur auf die Fahrradnutzung.

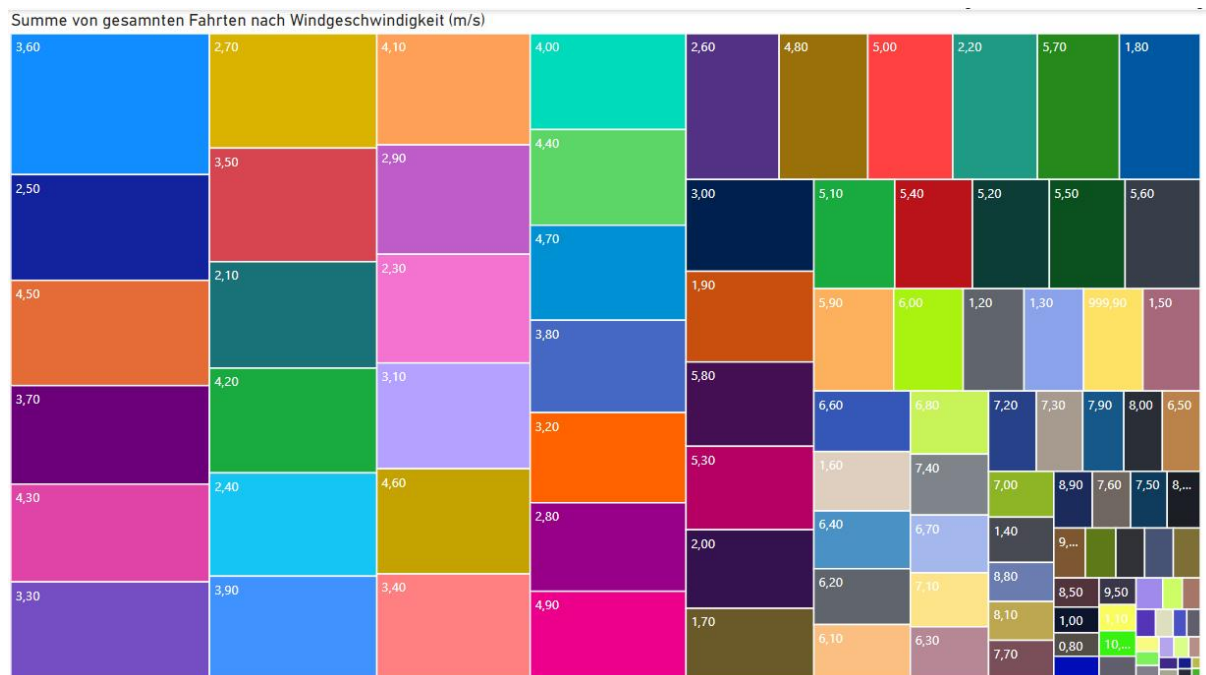


Abbildung 33 : Beziehung zwischen durchschnittlicher Windgeschwindigkeit (m/s) und Anzahl der Fahrten

(Quelle : Power BI Software – eigene Darstellung)

Es fällt bei dieser Visualisierung auf, dass die höchste Anzahl an Fahrten bei einer Windgeschwindigkeit von 3,6 m/s verzeichnet wurde. Diese Geschwindigkeit entspricht dem Punkt, an dem Blätter und dünnere Zweige in Bewegung geraten und Wimpel sich strecken.

Die Verwendung von Kästen zur Visualisierung der Daten stellt sie auf anschauliche Weise dar: Je größer der Kasten in den Zellen ist, desto mehr Fahrten wurden in Verbindung mit der spezifischen Windgeschwindigkeit verzeichnet.

Diese visuelle Darstellung ermöglicht eine schnelle Identifizierung von Bereichen mit besonders hoher oder niedriger Fahrradaktivität in Abhängigkeit von der Windgeschwindigkeit.

Die klare und übersichtliche Aufbereitung der Korrelationsdaten in Form von Kästen erleichtert die Interpretation und die Identifizierung von Mustern, was eine solide Grundlage für weiterführende Analysen und strategische Entscheidungen im Bereich städtischer Mobilität schafft.

4.2.4.2.2 Datenvisualisierung-/Analyse für das Sommer-Saison 2022

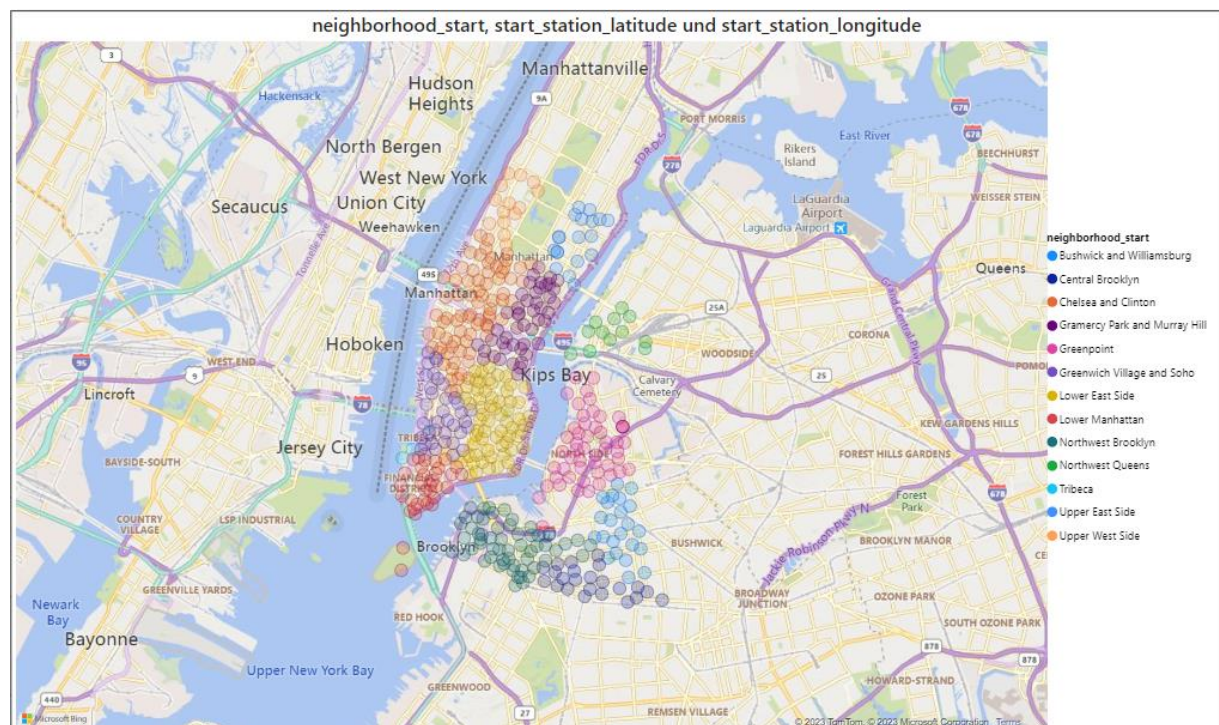


Abbildung 34 : Darstellung der Startstationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken

(Quelle : Power BI Software – eigene Darstellung)

Die geografische Visualisierung der Fahrradnutzung in New York City innerhalb eines Stadtgebiets eröffnet einen äußerst effektiven Ansatz zur Identifizierung geografischer Muster und Hotspots in Bezug auf die Verwendung von Fahrrädern. In dieser visuellen Darstellung erscheinen die Punkte in lebhaften Farben und repräsentieren verschiedene Stadtgebiete. Die Größe der Punkte ist standardisiert, da jeder Punkt eine Fahrradstation darstellt, die als Ausgangspunkt für Fahrradfahrten dient. Diese Standardgröße ist sinnvoll, da sie unabhängig von der jeweiligen Station ist und somit die gleichmäßige Darstellung gewährleistet.

Jede dieser Stationen stellt einen Startpunkt für Fahrradfahrten dar, und für einige dieser Stationen liegen noch keine zeitlichen Aufzeichnungen vor. Die Visualisierung ermöglicht es, die geografischen Verteilungsmuster der Fahrradnutzung in der Stadt leicht zu erfassen und Hotspots oder Bereiche mit besonders hoher Aktivität zu identifizieren. Dies kann wertvolle Erkenntnisse für die Verbesserung des Fahrradverkehrs in New York City liefern, da sie aufzeigt, welche Stadtteile besonders beliebt für Fahrradfahrten sind und wo möglicherweise zusätzliche Ressourcen für Fahrradstationen oder Infrastruktur benötigt werden.

In der nächsten Abbildung wird die Skalierung der Kreise in Abhängigkeit von der Dauer der Fahrt an jeder Station visualisiert. Dies ermöglicht es, visuell darzustellen, wie die Nutzungsdauer von Fahrrädern an verschiedenen Standorten variiert. Je größer der Kreis ist, desto länger sind die durchschnittlichen Fahrradfahrten an den jeweiligen Stationen.

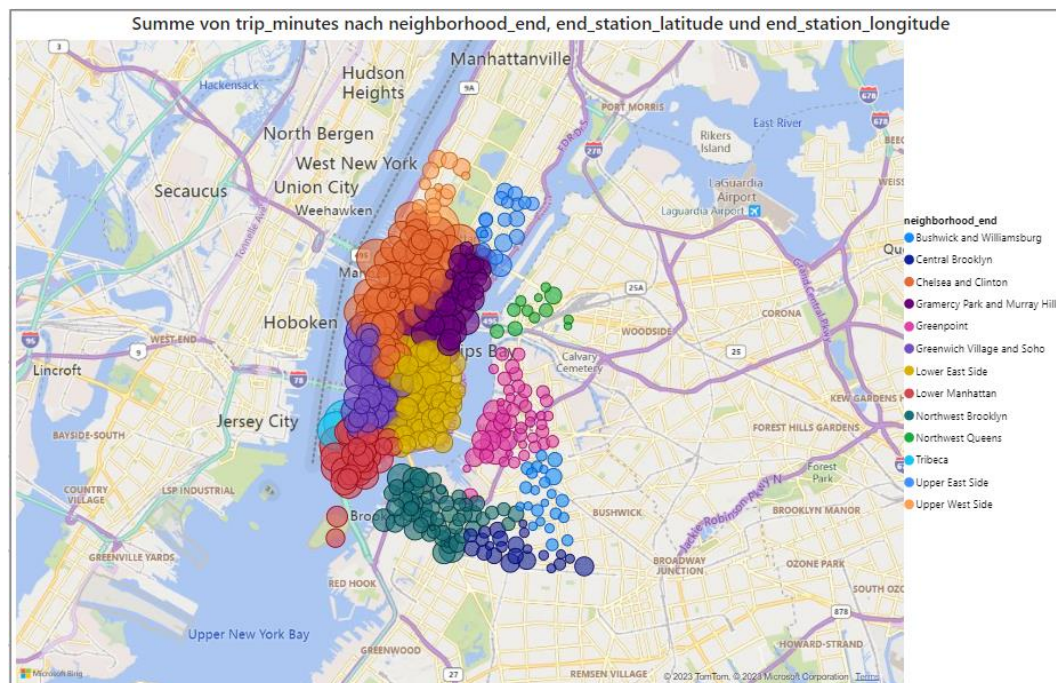


Abbildung 35 : Darstellung der Abgabestationen für Fahrradausleihen unterteilt nach Stadtteilen und Bezirken

(Quelle : Power BI Software – eigene Darstellung)

Diese visuelle Abbildung dient zur Darstellung der Fahrradnutzung innerhalb der New York Stadt und gewährt Einblicke in die Endstationen von Fahrradausleihen, welche nach Stadtteilen und Bezirken aufgeschlüsselt sind. Sie ermöglicht die Identifikation räumlicher Muster und Trends in Bezug auf die Fahrradnutzung. Die Hauptmerkmale dieser grafischen Darstellung umfassen:

Farbliche Codierung: Jeder Stadtteil oder Bezirk wird durch eine eindeutige Farbe repräsentiert, was eine leicht verständliche Identifikation ermöglicht. Die Farben auf der Abbildung entsprechen den verschiedenen geografischen Bereichen.

Kreise für Rückgaben: Die Endstationen sind durch Kreise gekennzeichnet, wobei die Größe dieser Kreise proportional zur Anzahl der zurückgegebenen Fahrräder ist. Größere Kreise weisen auf stark frequentierte Endstationen hin, während kleinere Kreise auf weniger frequentierte Stationen hinweisen.

Interaktivität: Die Abbildung kann interaktiv gestaltet sein, wodurch Nutzer in der Lage sind, mit der grafischen Darstellung zu interagieren. Ein Mausklick auf einen Kreis ermöglicht den Zugriff auf detaillierte Informationen zur entsprechenden Endstation.

Die bereitgestellte Abbildung gewährt einen Überblick über die räumliche Verteilung der Endstationen für Fahrradausleihen und ist von großer Bedeutung für urbane Planung, Verkehrsmanagement und die Optimierung von Fahrradverleihsystemen. Benutzer können rasch identifizieren, in welchen Stadtteilen die meisten Rückgaben erfolgen und wie sich die Rückgaben auf die verschiedenen Bezirke verteilen.

Kapitel 5 : Auswertung der Visualisierungstechnologien und Ergebnis der Forschung

In diesem Kapitel werden Ergebnisse der Expertise erörtert und Schlüsse aus dem praktischen Teil gezogen. Darüber hinaus wird eine Auswertung zwischen den im Rahmen der Arbeit verwendeten Technologien durchgeführt.

5.1. Auswertung der Ergebnisse aus den Visualisierungen & Antworten auf Geschäftsfragen

1) Verstehen, was Kunden wollen, was ein erfolgreiches Produkt ausmacht und wie neue Stationen die Nachfrage in verschiedenen geografischen Gebieten lindern könnten.

Die Forschung hat aufschlussreiche Erkenntnisse darüber geliefert, wie Kunden Fahrräder nutzen und welche Bedürfnisse sie dabei haben. Es wurde festgestellt, dass Fahrräder hauptsächlich für kurzfristige Ziele und Aktivitäten genutzt werden. Die Nutzung von Fahrrädern konzentriert sich vor allem auf zentrale Lagepunkte. Dies legt nahe, dass die meisten Menschen Fahrräder als eine Form des Kurzstreckentransports betrachten.

Ein erfolgreiches Produkt in diesem Kontext könnte daher darin bestehen, die Wahrnehmung von Fahrrädern zu verändern. Dies könnte durch gezieltes Marketing erreicht werden, das die Menschen dazu ermutigt, Fahrräder nicht nur als Transportmittel für kurze Strecken, sondern auch als Mittel zur Förderung von Sport und Bewegung zu betrachten.

Dies könnte beispielsweise durch Partnerschaften mit lokalen Unternehmen oder Stadtverwaltungen erreicht werden, um das Fahrradfahren als Sport und Freizeitaktivität zu fördern. Ebenso könnten Pilotprojekte in geografischen Gebieten mit hoher Nachfrage gestartet werden, um das Interesse an Fahrrädern für längere Strecken zu wecken.

Zusätzlich könnten Anreize wie Rabatte oder Sonderangebote in Gebieten mit niedriger oder keiner Nachfrage eingeführt werden, um die Nutzung von Fahrrädern in diesen Gebieten zu steigern. Die genannten Gebiete wie Bushwick and Williamsburg, Kingsbridge & Riverdale, Sunset Park, Long Island City, Morningside Heights, Central Harlem und Sunset Park könnten von gezielten Maßnahmen profitieren, um die Nachfrage anzukurbeln.

Die Einführung neuer Stationen kann ebenfalls dazu beitragen, die Nachfrage in verschiedenen geografischen Gebieten zu lindern. Dies sollte auf einer gründlichen Bedarfsanalyse basieren, um geografische Gebiete zu identifizieren, in denen die Nachfrage nach Fahrrädern hoch ist, aber bisher nicht ausreichend bedient wird. Durch die strategische Platzierung neuer Stationen in diesen Gebieten können Sie dazu beitragen, den Bedarf zu decken und die Verfügbarkeit von Fahrrädern zu erhöhen.

Insgesamt zeigt die Forschung, dass es vielfältige Möglichkeiten gibt, um die Nachfrage nach Fahrrädern in verschiedenen geografischen Gebieten zu steigern und die Nutzung von Fahrrädern für verschiedene Zwecke zu fördern.

Dies erfordert eine gezielte Herangehensweise, die sowohl Marketing- als auch Infrastrukturmaßnahmen umfasst.

2) Verstehen, wie das aktuelle Fahrradsortiment genutzt wird.

Die Analyse der Nutzung des aktuellen Fahrradsortiments hat interessante Einblicke ergeben. Die Mehrheit der Nutzer besteht hauptsächlich aus Männern im mittleren Alter, durchschnittlich etwa 38 Jahre alt. Diese Gruppe dominierender Nutzer bevorzugt das Fahrradverleihsystem, um ihre Fortbewegung zu gestalten.

Die beliebtesten Zielstationen konzentrieren sich auf zentrale Lagepunkte in der Stadt. Hierzu gehören Stadtviertel wie Chelsea und Clinton, Gramercy Park und Murray Hill, Lower East Side, Tribeca, Lower Manhattan sowie Greenwich Village und Soho. Diese Gebiete scheinen besonders attraktiv für Fahrradnutzer zu sein und zeichnen sich durch eine hohe Nachfrage aus.

Im Jahr 2022 wurde ein bemerkenswerter Rekord aufgestellt, da in diesem Jahr die höchste Anzahl von Fahrradfahrten registriert wurde. Insgesamt wurden 13.354.574 Fahrten verzeichnet, was auf ein starkes Interesse und eine steigende Akzeptanz des Fahrradverleihsystems in der New York Stadt hinweist.

Darüber hinaus wurden bestimmte Monate als besonders beliebt für Fahrradfahrten identifiziert. Diese Monate sind Mai, Juli, August und September, in denen die Nachfrage nach Fahrrädern deutlich ansteigt. Diese Erkenntnisse sind von unschätzbarem Wert, um die Dienstleistungen des Unternehmens zu optimieren und den Bedürfnissen unserer Kunden besser gerecht zu werden.

3) Wie kann man die Erkenntnisse über die Kundennutzung nutzen, um neue Stationen zu entwickeln?

Die Erkenntnisse über die Kundennutzung bieten eine wertvolle Grundlage, um gezielt neue Stationen für das Fahrradverleihsystem zu entwickeln und so die Bedürfnisse der Kunden noch besser zu erfüllen. Basierend auf den bereits gewonnenen Erkenntnissen eröffnen sich verschiedene Möglichkeiten:

1. Bedarfsermittlung in unterversorgten Gebieten: Man kann die Daten nutzen, um geografische Bereiche zu identifizieren, in denen die Nachfrage nach dem Service hoch ist, aber bisher unzureichend bedient wird. Die Einrichtung neuer Stationen in diesen Gebieten kann dazu beitragen, diese Lücke zu schließen und den Kunden besseren Zugang zu den Fahrrädern zu bieten.

2. Optimierung der Standortwahl: Die bisherigen Erkenntnisse über die beliebtesten Zielstationen und die häufigsten Nutzungszeiten ermöglichen es, neuen Stationen strategisch zu platzieren. Man kann somit sicherstellen, dass sie in Gebieten mit hoher Nachfrage und potenziell starkem Kundeninteresse eingerichtet werden.

3. Angepasste Angebote für Kundensegmente: Die Segmentierung der Kundenbasis nach Nutzungsverhalten ermöglicht es, maßgeschneiderte Angebote für verschiedene Kundengruppen zu entwickeln. Man kann beispielsweise spezielle Dienstleistungen oder Fahrradtypen für Pendler, Freizeitnutzer oder andere Zielgruppen bereitstellen.

4. Saisonal angepasste Stationen: Die Erkenntnis, dass bestimmte Monate besonders beliebt für Fahrradfahrten sind, erlaubt es, temporäre Stationen oder spezielle Angebote während dieser Saisons zu entwickeln. Dies kann die Nutzung während der Spitzenzeiten weiter steigern.

5. Feedback in die Entwicklung einfließen lassen: Das kontinuierliche Sammeln von Kundenfeedback und die Analyse von Kundenzufriedenheitsbewertungen ermöglichen es, neue Stationen und Dienstleistungen auf Basis der Wünsche und Anregungen der Kunden zu gestalten.

6. Partnerschaften zur Standortentwicklung: Die Daten können auch dazu verwendet werden, potenzielle Partnerschaften mit Unternehmen, Stadtverwaltungen oder anderen relevanten Akteuren zur Standortentwicklung zu erkunden. Diese Partner können sowohl Standorte als auch Finanzierung bereitstellen.

Durch die gezielte Anwendung dieser Erkenntnisse kann man sicherstellen, dass die Entwicklung neuer Stationen nicht nur den Bedürfnissen der Kunden entspricht, sondern auch zur Steigerung der Kundenzufriedenheit und zur Förderung des umweltfreundlichen Verkehrs beiträgt. Die Daten sind somit ein unschätzbares Werkzeug, um die Dienstleistungen zu optimieren und weiterhin auf die Anforderungen der Nutzer einzugehen.

4) Wie verschiedene Nutzer (Abonnenten und Nicht-Abonnenten) die Fahrräder benutzen. Es wird eine große Gruppe von Nutzern untersucht, um eine faire Repräsentation der Nutzer an verschiedenen Standorten und mit niedrigen bis hohen Aktivitätsniveaus zu erhalten.

Im analysierten Zeitraum ergibt sich, dass von der Gesamtzahl von 51.228.879 Nutzern 45.803.961 als "Subscriber" klassifiziert sind, was den überwiegenden Teil der Gesamtnutzer ausmacht. Diese hohe Anzahl von Nutzern deutet darauf hin, dass eine bedeutende Anzahl von Fahrradservice-Nutzern regelmäßige Abonnenten ist, die den Service in wiederholter Form nutzen.

Hingegen zeigt das Diagramm, dass 5.424.918 Nutzer als "Customer" eingestuft sind. In diesem Kontext könnten "Customers" gelegentliche oder erstmalige Nutzer sein, die den Service nicht in regelmäßigen Abständen abonnieren, sondern ihn je nach Bedarf nutzen.

5) Wie beeinflusst das Wetter die Nutzung von Fahrrädern?

Die Analyse des Einflusses des Wetters auf die Fahrtenhäufigkeit hat klare Trends und Muster aufgedeckt. Es hat sich gezeigt, dass schlechtes Wetter einen negativen Einfluss auf die Fahrradnutzung hat. Insbesondere bei einer Niederschlagswahrscheinlichkeit von 0% wurde die höchste Anzahl von Fahrten registriert. Je höher die Niederschlagswahrscheinlichkeit ansteigt, desto weniger Fahrten wurden verzeichnet. Dies legt nahe, dass Regen und schlechtes Wetter die Bereitschaft der Menschen, Fahrräder zu nutzen, erheblich verringern.

Auch die Temperatur hat einen signifikanten Einfluss auf die Fahrtenhäufigkeit. Die meisten Fahrten wurden bei einer Temperatur von 38 Grad Celsius registriert. Nach dieser Schwelle sank die Anzahl der Fahrten, was darauf hinweist, dass extrem hohe oder niedrige Temperaturen die Nutzung des Fahrradverleihsystems beeinträchtigen.

Hinsichtlich der Windgeschwindigkeit konnte festgestellt werden, dass Werte zwischen 1,7 m/s und 5,6 m/s eine konsistente Fahrtenhäufigkeit aufwiesen. Sobald die Windgeschwindigkeit über 5,6 m/s stieg, nahm die Anzahl der Fahrten drastisch ab. Dies zeigt, dass starke Winde eine entscheidende Rolle bei der Entscheidung der Menschen spielen, ob sie das Fahrrad nutzen oder nicht.

Insgesamt unterstreicht die Analyse des Wettereinflusses auf die Fahrtenhäufigkeit die Notwendigkeit, bei der Planung und Verwaltung von Fahrradverleihsystemen die Wetterbedingungen zu berücksichtigen. Dies kann die Grundlage für gezielte Marketingkampagnen, saisonale Anpassungen und die Verbesserung der Servicebereitstellung sein, um die Auswirkungen ungünstiger Wetterbedingungen auf die Fahrradnutzung zu mildern.

5.2. Auswertung beider Visualisierungstechnologien : Power BI und Tableau

5.2.1. Power BI

5.2.1.1 Stärken :

1. Microsoft investiert kontinuierlich in die Verbesserung von BI-Diensten seiner Hauptprodukte: Microsoft Office, Microsoft SharePoint und Microsoft SQL Server. Dadurch steigert das Unternehmen den Wert und macht die Nutzung seiner Plattform für Entwickler attraktiver. Somit bietet das Unternehmen eine Vielzahl wettbewerbsfähiger BI-Dienste, Pakete und Preise und stellt sicher, dass die Akzeptanz insbesondere in Unternehmen, die eine Microsoft-zentrierte Infrastruktur nutzen, hoch bleibt.⁶¹ Darum verzeichnet Microsoft einen erheblichen Marktanteil und eine beeindruckende Zahl treuer Benutzer

2. Außerdem zeichnet sich Microsoft durch das geschickte Bündeln von Kosten und BI-Leistungen aus, was das Preis-Leistungs-Verhältnis des Unternehmens stärkt. Es werden kostengünstige Preise im Vergleich zu den kommerziellen Konkurrenten angeboten. Des Weiteren baut Microsoft seine BI-Funktionen in Produkte aus, die die meisten Unternehmen bereits besitzen, wie SQL-Server, Microsoft Office oder SharePoint. Das macht es für viele Unternehmen schwierig, Alternativen zu rechtfertigen.⁶²

3. Darüber hinaus wird dem Microsoft Excel große Aufmerksamkeit geschenkt. Microsoft arbeitet daran, die Berichts- und Datenermittlungsfunktionen zu verbessern, was dazu führt, dass die Analysewerkzeuge innerhalb von Excel nicht nur für Entwickler, sondern auch für nicht IT-affine Mitarbeiter einfacher werden. Somit können maximal viele Personen von deren Funktionen profitieren, von Sekretären bis zu Managern und Geschäftsführern. Die Veröffentlichung von SQL Server Power Pivot und Power View hat dazu beigetragen, die Interaktivität im Vergleich zu traditionellen berichtsorientierten Architekturen zu erhöhen.⁶³

⁶¹(vgl. Hagerty et al. 2012, S. 15-16)

⁶² (vgl. ebd.: S. 15-16)

⁶³ (vgl. Ferrari/Russo 2018, S. 13)

4. Die IT-Abteilung erhält durch die Überwachungsfunktionen von Power Pivot-Arbeitsmappen in SharePoint eine verbesserte Kontrolle über freigegebene Inhalte und die Validierung von Datenquellen und Modellen innerhalb dieser Arbeitsmappen im Vergleich zu anderen eigenständigen Datenermittlungsprodukten. Microsoft plant, diese verwalteten Funktionen für Geschäftsanwender mit der Veröffentlichung von SQL-Server 2012 weiter auszubauen. Diese Erweiterungen sollen es Power Pivot ermöglichen, reibungslos von einer individuellen Arbeitsmappe zu einer unternehmensweiten Datenquelle und Bereitstellung zu wechseln. Dabei sollen Geschäftsanwender in der Lage sein, die Kontrolle über den gesamten Prozess zu behalten und gleichzeitig die nahtlose Integration von persönlichen und unternehmensweiten Daten zu gewährleisten.⁶⁴ Diese Fähigkeit kann dazu beitragen, die Unterschiede zwischen Unternehmen und Abteilungssilos zu überbrücken, ohne dass die Geschäftsanwender beeinträchtigt werden.⁶⁵

Aufgrund der begrenzten Fähigkeit von Excel, bis zu einer Million Zeilen zu laden, ermöglicht der Power Pivot Ansatz eine anspruchsvolle Analyse bei umfangreichen Datenmengen. Die Höchstgrenze für die Anzahl der Zeilen, die in eine Power Pivot Tabelle geladen werden können, besteht tatsächlich nicht.⁶⁶

5. Microsoft OLAP-Funktionen, laut Duncan/Mabee, zeigen sich als die beliebteste Auswahl bei Unternehmen. Dabei spielen Akzeptanz, Microsoft Frontend-Tools und die Funktionalität von SQL Server Analysis Services eine bedeutende Rolle.⁶⁷ Um den Anforderungen von Unternehmen gerecht zu werden, die auf neuere In-Memory-OLAP-Architekturen statt auf herkömmliche OLAP-Architekturen angewiesen sind, wird Microsoft in SQL-Server 2012 eine vollständige In-Memory-Version von Microsoft Analysis Services-Cubes einführen. Dies basiert auf den In-Memory-Fähigkeiten von SQL Server Power Pivot. In-Memory-Datenbanksysteme verbessern die Leistung, indem sie die erweiterten Arbeitsspeicherkapazitäten moderner Datenbanksysteme nutzen. Es ist möglich, sowohl relationale als auch nicht-relationale In-Memory-Datenbanken zu verwenden.⁶⁸

Zu den wesentlichen Stärken von In-Memory-Datenbanksystemen zählt die extrem schnelle Datenzugriffsgeschwindigkeit im Arbeitsspeicher im Vergleich zur herkömmlichen Datenträgersubsystemen. In vielen SQL-Server-Anwendungsfällen passt der gesamte Arbeitssatz in den verfügbaren Arbeitsspeicher, was eine spürbare Leistungssteigerung ermöglicht.

Die Einführung des Persistierten Protokollpuffers im Service Pack 1 von SQL-Server 2016 hat schreibintensive Workloads erheblich verbessert.

⁶⁴ (vgl. Hagerty et al. 2012, S. 16)

⁶⁵ (vgl. Ferrari/Russo 2018, S. 24)

⁶⁶ (vgl. ebd.: S. 22)

⁶⁷ (vgl. Duncan/Mabee 2023)

⁶⁸ (vgl. Hagerty et al. 2012, S. 17)

Mit dieser Funktion wird die Anzahl der für das Festschreiben des Protokolls erforderlichen Vorgänge reduziert, was die Gesamtleistung und Effizienz steigern. Diese Fortschritte ermöglichen es, traditionelle Konzepte zu überwinden und die Leistung von SQL-Server weiter zu optimieren.⁶⁹

5.2.1.2 Schwächen :

1. Die vom Gartner durchgeführte Umfrage an BI-Plattform-Kunden hat gezeigt, dass im gleichen Jahr der Veröffentlichung dieses Berichts (2012) bestimmte Kriterien, wie Produktfunktionalität und Kundensupport, zurückgegangen sind. Somit befanden sich diese Kriterien unter dem Durchschnitt.⁷⁰ Kundensupport ist ein wichtiger Baustein für die Kundenorientierung, wie bereits zu Beginn dieser Arbeit erwähnt wurde. Somit beeinflusst dieses explizit den Erfolg des Produktes.

2. Wenn die Produkte an sich komplex sind, stellt dies eine große Herausforderung dar. Da die BI-Plattform von Microsoft aus drei verschiedenen Tools (Office, SQL-Server und SharePoint) besteht, die auch Nicht-BI-Funktionen ausführen, ist die Integration der erforderlichen Komponenten und die Erstellung der Anwendungen dem Unternehmen überlassen. Durch den Do-it-yourself-Ansatz von Microsoft liegt ein größerer Teil der Entwicklung von BI-Lösungen und der Integration der Plattformkomponenten bei den Kunden, verglichen mit den All-in-One-BI-Plattformen, die von den meisten anderen Anbietern auf dem BI-Markt angeboten werden. Microsofts Roadmap für Office, die die Konsolidierung von mehr und mehr Front-End-Berichts-, Dashboard- und Analysefunktionen in Excel vorsieht, sollte im Laufe der Zeit einen Teil dieser Komplexität beseitigen. Obwohl BI in der Cloud für die meisten Unternehmen in der Magic-Quadrant-Umfrage noch keine hohe Priorität hat, hat Microsoft die Cloud-Bereitstellung ganz oben auf seine Liste der wichtigsten Entwicklungs- und Markteinführungsinitiativen für BI gesetzt, indem es seine Kern-BI-Produkte - SQL Server, SharePoint und Office - schließlich in der Cloud verfügbar macht. Diese Investition und Betonung ist ein Kernstück der Microsoft-Strategie, BI einfach und kostengünstig bereitzustellen - Cloud-basierte BI wird theoretisch einen Teil der Komplexität der Drei-Komponenten-Anforderung beseitigen.⁷¹

⁶⁹ (vgl. Michael Carrig/Assaf 2023)

⁷⁰ (vgl. Hagerty et al. 2012, S. 17-18)

⁷¹ (vgl. ebd.: S. 17-18)

5.2.2. Tableau

5.2.2.1 Stärken:

1. Tableau bietet leistungsstarke, interaktive Visualisierungen für Analysen, Berichte und Informationsbereitstellung. Somit gilt es als eine der attraktivsten Analyse-Softwares im Bereich Business Intelligence (BI). In einem wettbewerbsintensiven Markt zeichnet sich Tableau nicht nur durch seine Leistungsstärke, sondern auch durch sein benutzerfreundliches und intuitives Interface aus. Die Bereitstellung der BI-Tools erfolgt einfach und ohne IT-Unterstützung.⁷²

2. Dank der spezifischen geschäftsorientierten Daten-Mashup-ETL-Funktionen und Schnittstellen für Daten aus verschiedenen Ressourcen, die die Tableau-eigene Technologie namens VizQL nutzen.⁷³

Diese Technologie wird durch eine Drag-and-Drop-Operation von Daten ausgelöst, die dazu dient, eine SQL- oder MDX-Abfrage zu verstehen und zu interpretieren, und im Anschluss daraus eine Visualisierung zu erstellen. Darüber hinaus verfügt Tableau über die spaltenbasierte In-Memory-Daten-Engine, was eine zügige Ausführung ermöglicht, und die Komplexität der Abfragen verringert, besonders wenn es sich um multidimensionale Daten handelt.

3. Die programmierfreie Data-Mashup-Fähigkeit Hand in Hand mit In-Memory-Datenbank machen es leicht, die Big Data zu mischen und zu visualisieren.⁷⁴ Dabei werden die Beziehungen zwischen verschiedenen Datenquellen schnell erkannt und daraus ein Schema ähnlich dem BPM-Konzept gezeichnet. Somit haben Benutzer Möglichkeit, sich mit einer Vielzahl von Datenquellen zu verbinden und interaktive Dashboards zu erstellen. Außerdem können sie die Grafik nach verschiedenen Parametern filtern.

4. Laut dem Gartner-Bericht nimmt das Volumen der analysierten Daten von Tableau kontinuierlich zu, insbesondere wenn es mit Tools wie Hadoop kombiniert wird. Darüber hinaus wächst die Zahl der externen Anwender ständig, die über die Website auf das Tool zugreifen. Attraktive SaaS-Angebote des Unternehmens tragen dazu bei, das Interesse der Kunden, wie zum Beispiel CBS-Sports und anderen Nachrichten-, Regierungs- und Medienwebseiten, zu steigern.

⁷² (vgl. Hagerty et al. 2012, S. 31-33)

⁷³ (vgl. ebd.: S. 31-33)

⁷⁴ (vgl. ebd.: S. 31-33)

Diese Kunden nutzen eingebettete Tableau-Visualisierungen und teilen sie mit ihren Zielgruppen. Dieser Ansatz, bekannt als Go-to-Market, hat Tableau einen großen Erfolg gegenüber dem traditionellen Vertriebskanalansatz verschafft.⁷⁵

5.2.2.2. Schwächen:

1. Im Vergleich zu anderen BI-Plattformen fehlt Tableau an umfassender Funktionsvielfalt, wie prädiktive Analysen und Machine Learning. Das Unternehmen konzentriert sich daher ausschließlich auf die Datenanalyse und interaktive Datenvisualisierung. Daher erfüllen die Produkte des Unternehmens oft einen ungedeckten Bedarf in Unternehmen und werden häufig als Ergänzung zu einer bestimmten BI-Plattform eingesetzt. Das könnte die langfristigen Aussichten von Tableau gefährden.⁷⁶

2. Darüber hinaus stellt der Gartner-Bericht fest, dass die internationale Präsenz des Unternehmens sehr begrenzt ist. Es werden nur Englisch, Deutsch, Portugiesisch, Italienisch, Chinesisch, Koreanisch und Französisch unterstützt.⁷⁷

3. Obwohl Tableau bereits eine Reihe von ORM-Partnern hat, hält es an einer horizontalen Plattformstrategie fest und hat noch keine Pläne zur Entwicklung einer vertikalen Strategie für bestimmte Branchen wie Gesundheitswesen, Industrie oder Online-Handel, etc.⁷⁸

5.3. Wesentliche Unterschiede „Tableau“ vs „Power BI“ aus Benutzersicht

Power BI und Tableau sind zwei leistungsstarke Datenvisualisierungstools, die in der Geschäftswelt weit verbreitet sind. Aus eigener Beobachtung haben diese Tools einige signifikante Unterschiede in ihrer Herangehensweise und Funktionalität:

1. Dashboard-Orientierung vs. Visualisierung: Power BI ist stark Dashboard-orientiert. Das bedeutet, dass Benutzer zunächst in einer Umgebung arbeiten, in der sie Dashboards erstellen. Im Gegensatz dazu beginnt Tableau in der Regel mit der Erstellung von Visualisierungen, und die Erstellung eines Dashboards erfolgt in einem späteren Schritt.

2. Datenmanipulation und Sprache: Power BI verwendet eine spezielle Sprache namens DAX (Data Analysis Expression), die es den Benutzern ermöglicht, Daten und Spalten direkt innerhalb der Power BI-Umgebung zu manipulieren.

⁷⁵ (vgl. ebd.: S. 31-33)

⁷⁶ (vgl. Hagerty et al. 2012, S. 31-33)

⁷⁷ (vgl. ebd.: S. 31-33)

⁷⁸ (vgl. ebd.: S. 31-33)

Dies bedeutet, dass Sie Berechnungen und Transformationen direkt in Power BI durchführen können, ohne einen Schritt zurück zur ETL (Extrahieren, Transformieren, Laden) -Prozess durchführen zu müssen. Tableau bietet auch Datenmanipulation, erfordert jedoch möglicherweise mehr Arbeit außerhalb der Anwendung, um umfassende Transformationen durchzuführen.

3. Visualisierungsbibliothek: Power BI enthält einen integrierten Visual Store, in dem Benutzer auf eine Vielzahl von Visualisierungen und Erweiterungen zugreifen können. Viele dieser Visualisierungen können kostenlos importiert werden und ermöglichen es den Benutzern, ihre Dashboards und Berichte mit verschiedenen Diagrammtypen und visuellen Elementen zu erweitern. Tableau bietet ebenfalls eine breite Palette von Visualisierungsoptionen, jedoch ist die Erweiterung der Visualisierungsfunktionalität möglicherweise nicht so nahtlos integriert wie in Power BI.

4. Integration: Power BI ist eng in die Microsoft-Produktfamilie integriert und bietet eine nahtlose Integration mit anderen Microsoft-Tools wie Excel und Azure. Tableau ist zwar vielseitiger in Bezug auf Datenquellen und Plattformen, erfordert jedoch möglicherweise zusätzliche Integrationsarbeit.

5.4. Ergebnis der Forschung

Die Anwendung des Business Intelligence Ansatzes mit Microsoft Azure und Power BI führt zu signifikanten Verbesserungen im Verständnis der Fahrradnutzung im Vergleich zur Verwendung von Google-Technologie in Kombination mit Tableau für die Analyse von „Cyclistic“ Unternehmensdaten in der Stadt New York. Die Untersuchung anhand von realen Unternehmensdaten und die Analyse mittels Power BI und Tableau haben gezeigt, dass die Nutzer der Fahrräder detaillierte Informationen liefern. Dies beinhaltet präzise Zahlen, Metriken und sogar Geodaten, die wesentlich zur Entscheidungsfindung beitragen. Das bessere Verständnis über das Geschäft, das durch die Anwendung des BI-Ansatzes erreicht wurde, unterstreicht die Effektivität von Microsoft Azure und Power BI für die Fahrradbranche.

Zusammenfassend lässt sich sagen, dass die vorliegende Bachelorarbeit erfolgreich die Forschungsfrage beantwortet hat. Die Ergebnisse zeigen konkrete visualisierte Daten und daraus gewonnene Erkenntnisse, was darauf hindeutet, dass der Ansatz unverzichtbar für jedes Geschäft sein könne. Diese Erkenntnisse tragen dazu bei, das Verständnis vom Geschäft zu vertiefen und könnten Implikationen für die Entscheidungsfindung und Unternehmensstrategie haben.

5.5. Handlungsempfehlungen

"Daten sind das Gold des 21. Jahrhunderts", dieser Spruch ist gerade in aller Munde. Für das Unternehmen ist es von hoher Bedeutung, jede Art der Geschäftsdaten in sauberer und konformer Form zu speichern und zu bewahren. Dies hat eine strategische Bedeutung für das Geschäft. Denn anhand dieser Daten kann nicht nur die aktuelle Lage beleuchtet werden, sondern auch zukünftige Prognosen mithilfe fortgeschrittener Ansätze wie KI und Machine Learning erstellt werden. Data Science ist eine umfangreiche Lehre, die ausschließlich auf Daten ausgerichtet ist. Dabei werden die Daten nach Beziehungen untersucht, und daraus werden spezielle Modelle erstellt und trainiert.

Gerade für kleine Unternehmen empfiehlt sich die Verwendung der Google Technologie. Diese ist maximal vereinfacht und verfügt über ein benutzerfreundliches Interface. Außerdem können dort fortgeschrittene Ansätze für Daten wie Künstliche Intelligenz und Machine Learning verwendet werden. Für größere Organisationen ist die Microsoft Azure Technologie eher geeignet. Diese ist deutlich komplexer und erfordert IT-affine Personen, wie beispielsweise Entwickler, die mit Programmiersprachen vertraut sind. Nichtsdestotrotz, um den vollen Nutzen aus den Daten zu ziehen, empfiehlt es sich, die Förderung einer datengetriebenen Kultur im gesamten Unternehmen durch verschiedene Schulungen, das Steigern der Datenkompetenz bei den Mitarbeitern, die Formulierung von Unternehmenszielen im Umgang mit Daten und ständige Überwachung sowie die Erstellung von Feedback-Schleifen, die die Funktionalität und Effizienz des BI-Ansatzes überwachen, um kontinuierliche Anpassung zu gewährleisten und auf den neusten Stand zu bleiben.

5.6. Fazit

In dieser Arbeit wird eingehend mit dem Thema Business Intelligence (BI) auseinandergesetzt. Von der Entwicklung einer BI-Strategie über die Architektur bis hin zu den Zielen und Technologien wurde jedes Detail gründlich behandelt. Der Fokus lag darauf, ein umfassendes Verständnis für die verschiedenen Aspekte von BI zu vermitteln. Durch die Analyse von Strategien und Technologien wird die Bedeutung einer effektiven BI-Implementierung klar, wodurch Unternehmen besser informierte Entscheidungen treffen können.

Im Rahmen dieser Arbeit wurde erfolgreich aufgezeigt, wie aus Millionen von unstrukturierten Rohdaten eine klare Übersicht erstellt werden kann. Dies ermöglicht nicht nur die Darstellung von Daten, sondern auch die Gewinnung wertvoller Einblicke, die ein generelles Verständnis für das Geschäft bieten.

Die in dieser Arbeit vorgestellte Methodik und Umsetzung kann als wegweisend betrachtet werden, da sie verdeutlicht, wie Unternehmen in der heutigen Ära, in der massive Datenmengen erzeugt werden, von solchen Ansätzen profitieren können. Business Intelligence (BI) stellt dabei nur den Anfang dar. Es ermöglicht Unternehmen, ihre Daten effizient zu organisieren und zu visualisieren. Dies schafft eine Grundlage für datengesteuerte Entscheidungsfindung.

Jedoch ist es wichtig zu betonen, dass BI nicht das Ende des Datenmanagements ist. Es gibt weitere Aspekte, die in der heutigen datengetriebenen Welt von Bedeutung sind. Data Science beispielsweise geht über BI hinaus und ermöglicht die Bildung konkreter Vorhersagen und die Identifikation von Mustern und Trends. Machine Learning und Künstliche Intelligenz (KI) bieten fortgeschrittene Möglichkeiten, um die Daten noch tiefergehend zu analysieren und automatisierte Entscheidungsprozesse zu entwickeln.

Insgesamt zeigt diese Arbeit auf, wie die Umsetzung einer Business Intelligence-Lösung als Demonstrator nicht nur für die Lehre, sondern auch für die Praxis von großer Relevanz ist. Sie verdeutlicht, wie Daten in wertvolles Wissen umgewandelt werden können, und weist gleichzeitig auf die vielfältigen Möglichkeiten hin, die sich in der Welt der Datenanalyse und -verarbeitung eröffnen. Damit leistet die Arbeit einen wichtigen Beitrag zur Förderung des Verständnisses und der Anwendung von datenbasierten Technologien in verschiedenen Anwendungsbereichen.

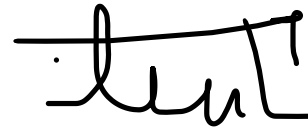
Es war eine faszinierende Reise durch die Welt der Business Intelligence, und ich hoffe, dass diese Arbeit einen Beitrag zum Verständnis und zur Wertschätzung dieses wichtigen Geschäftsfelds leisten konnte.

IV. Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen sind, sind als solche kenntlich gemacht. Die Arbeit hat in gleicher Form noch keiner anderen Prüfbehörde vorgelegen.

Frankfurt am Main, 21.11.2023

(Ort, Datum)

A handwritten signature in black ink, consisting of a horizontal line followed by a series of loops and a vertical stroke.

(Unterschrift)

V. Quellenverzeichnis

- [Sinz/Ulbrich-vom-Ende 2009] Sinz, E. J. ; Ulbrich-vom-Ende, A.: *Architektur von Data-Warehouse-Systemen*. In: Chamoni, P.; Gluchowski, P.: *Analytische Informationssysteme – Business Intelligence Technologien und – Anwendungen*. 4. Auflage, Springer- Verlag. Heidelberg, 2009, S. 175-196.
- Arul, Shankar (2021). "Tableau for Business Users.", [online]
<https://link.springer.com/book/10.1007/978-1-4842-7786-7>
- Azvine, B./Zhan Cui/Detlef Nauck (2005): Towards real-time business intelligence, in: *Bt Technology Journal*, Bd. 23, Nr. 3, S. 214–225, [online] doi:10.1007/s10550-005-0043-0.
- Bauer, Andreas (2009): *Data-Warehouse-Systeme: Architektur, Entwicklung, Anwendung*. Heidelberg
- Carlisle, Stephanie (2018): Software: Tableau and Microsoft Power BI, in: *Technology architecture + design*, Taylor & Francis, Bd. 2, Nr. 2, S. 256–259, [online]
doi:10.1080/24751448.2018.1497381.
- Chamoni, Peter/Peter Gluchowski (2010): *Analytische Informationssysteme: Business Intelligence-Technologien und -Anwendungen*, Springer.
- Dataflow – Übersicht (o. D.): in: *Google Cloud*, [online]
<https://cloud.google.com/dataflow/docs/overview?hl=de>.
- Duncan, Owen/Dan Mabee (2023): Microsoft - Erweiterte von OLAP-Funktionalität, learn.microsoft.com, [online] <https://learn.microsoft.com/de-de/analysis-services/multidimensional-models/extending-olap/extending-olap-functionality?view=asallproducts-allversions>.

- Ferrari, Alberto/Marco Russo (2018): *Datenanalyse mit Microsoft Power BI und Power Pivot für Excel*, dpunkt.verlag, [online] <https://d-nb.info/1151638951>.
- Fischer, Felix B./Anton A. Burger/Benedikt Gehling (2023): Aktuelle Herausforderungen bei der Implementierung von Self-Service Business Intelligence, HMD Praxis der Wirtschaftsinformatik. [online] : <https://link.springer.com/article/10.1365/s40702-023-00962-4>
- Franks, Bill (2014): *The Analytics Revolution*, John Wiley & Sons, Inc. eBooks, [online] doi:10.1002/9781118936672.
- Fröhlich, Martin/Kurt Glasner (2007): *IT Governance - Leitfaden für eine praxisgerechte Implementierung*, Wiesbaden.
- Gablerlexikon.2004. *Stichwort Berichtswesen*. Wiesbaden : Gabler
- Gansor, Tom/Andreas Totok/Steffen Stock (2010): *Von der Strategie zum Business Intelligence Competency Center (BICC)*, Carl Hanser Verlag GmbH & Co. KG eBooks, [online] doi:10.3139/9783446424869.
- Gelbrich, Katja/Stefan Wünschmann/Stefan Müller (2018): *Erfolgsfaktoren des Marketing*, Vahlen.
- Hagerty, John/Rita L. Sallam/James Richardson (2012): *Magic quadrant for Business Intelligence Plattformen*, Gartner, [online] https://www.existbi.com/wp-content/uploads/2013/09/18bf22fa59f21c06de6e4db354357322_196871334155831_GartnerMagicQuadrantforDataQualityTools2010.pdf.
- Hahne, Michael (2014): *Modellierung von Business-Intelligence-Systemen: Leitfaden für erfolgreiche Projekte auf Basis flexibler Data-Warehouse-Architekturen*, dpunkt.verlag.
- Hassanien, Aboul Ella/Dequan Zheng/Zhijie Zhao/Zhipeng Fan (2023): *Business Intelligence and Information Technology: Proceedings of BIIT 2022*, Springer Nature

Hoffmann, Stefan/Payam Akbar (2023): *Consumer Behavior: Understanding Consumers—Designing Marketing Activities*, Springer Nature.

Hungenberg, Harald (2004): *Strategisches Management in Unternehmen*, Wiesbaden.

Kari, Mohamed/Felix Weber/Reinhard Schütte (2019): Datengetriebene Entscheidungsfindung aus strategischer und operativer Perspektive im Handel, in: *HMD. Praxis der Wirtschaftsinformatik*, Springer Science+Business Media, Bd. 56, Nr. 5, S. 914–931, [online] doi:10.1365/s40702-019-00530-9.

Kemper, Hans-Georg/Henning Baars/Walid Mehanna (2010): *Business Intelligence - Grundlagen und praktische Anwendungen: Eine Einführung in die IT-basierte Managementunterstützung*, Springer Science & Business Media.

Kohne, Andreas (2018): *Cloud-Föderationen: SLA-basierte VM-Scheduling-Verfahren*, Springer-Verlag.

Meffert, Heribert (2019): *Marketing : Grundlagen marktorientierter Unternehmensführung : Konzepte - Instrumente - Praxisbeispiele ; mit neuer Fallstudie VW Golf*, Gabler eBooks.

Michael Carrig, Brian/William Assaf (2023): Microsoft - In-Memory-Datenbanksysteme und -Technologien, learn.microsoft.com, [online] <https://learn.microsoft.com/de-de/sql/relational-databases/in-memory-database?view=sql-server-ver16>.

Mihart (2023): What is Power BI? - Power BI, Microsoft Learn, [online] : <https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>.

Negash, S. (2004). *Business Intelligence. Communications of the Association for Information Systems*, 13, pp-pp. <https://doi.org/10.17705/1CAIS.01315>

Otto, Boris/Kristin Wende/Alexander Schmidt/Kai Hüner/Tobias Vogel (2008): Unternehmensweites Datenqualitätsmanagement: Ordnungsrahmen und Anwendungsbeispiele, in: *Springer eBooks*, S. 211–230, [online] doi:10.1007/978-3-540-77578-2_10.

- Poßner, Inken (2023): *Data governance in Germany - an introduction*, Kassel : Kassel university press.
- Reinking, Oliver (2023): *Business Intelligence - eine Einführung*.
- RICHARDSON, J., R. SALLAM, K. SCHLEGEL, A. KRONZ und J. SUN, 2020 : *Magic Quadrant for Analytics and BusinessIntelligence Platforms*[online]. 11. Februar 2020 [Zugriff am: 9. Juli 2020]. Verfügbar unter:
https://www.gartner.com/doc/reprints?id=1-3TXXSLV&ct=170221&st=sb&ocid=mkto_eml_EM597235A1LA1?ocid=eml_pg165511_gdc_comm_b%E2%80%A6
- Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr. (o. D.): [online] <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:de:html>.
- Scherrer, G. (1991): *Kostenrechnung*, Stuttgart : Lucius & Lucius.
- Schneider, Nico (2020): *Business Intelligence Entwurf und prototypische Implementierung einer Omnichannel-Marketinganalyseplattform zur kanalübergreifenden Kampagnen-Steuerung auf Basis von PowerBI*, Bachelorarbeit, Hochschule Esslingen of Applied Sciences.
- Schön, Dietmar (2023): *Planning and reporting in BI-supported controlling: Fundamentals, Business Intelligence, Mobile BI, Big Data Analytics and AI*, Springer Nature.
- Schrödl, Holger (2009): *Business Intelligence mit Microsoft SQL Server 2008*, Carl Hanser Verlag GmbH & Co. KG eBooks, [online] doi:10.3139/9783446420748.
- Schroiff, Hans-Willi/Tina Müller (2013): *Warum Produkte floppen*, [online] <https://content-select.com/en/portal/media/view/5885fab0-37ac-4153-b2a0-712fb0dd2d03>
- Seiter, Mischa (2023): *Business Analytics: wie Sie Daten für die Steuerung von Unternehmen nutzen*.

Trommsdorff, Volker/Thorsten Thorsten Teichert (2011): *Konsumentenverhalten*, 8. Aufl., Stuttgart, Kohlhammer Verlag.

Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG. (2016): [online]
<https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=celex%3A32016R0679>.

Was ist Azure Analysis Services? (o. D.): learn.microsoft.com, [online]
<https://learn.microsoft.com/de-de/azure/analysis-services/analysis-services-overview>.

Was ist BigQuery? (o. D.): in: *Google Cloud*, [online]
<https://cloud.google.com/bigquery/docs/introduction?hl=de>.

Wild, J. (1981): *Grundlagen der Unternehmensplanung*.