**Data Understanding**

Der Datensatz "Verkehrsunfälle in England" enthält insgesamt 21 Spalten und 307.974 Zeilen und stammt von Kaggle. Diese Daten lassen sich in vier Hauptaspekte unterteilen: Ort, Wetter, Zeit und Art des Unfalls. Unter dem Aspekt des Ortes werden für jeden Unfall die geografischen Koordinaten mit Breitengrad und Längengrad ("Longitude" und "Latitude") sowie die zuständige Polizeibehörde ("Police Force") und die örtliche Verwaltung ("Local Authority (District)") erfasst. Auch zur Urbanität oder Rurität ("Urban\_or\_Rural\_Area") des Unfallortes werden Angaben gemacht. Zeitlich wird sowohl Tag des Unfalls ("Accident Date"), der Wochentag ("Day\_of\_week") als auch die Zeit ("Time") aufgezeichnet. Die Wetterbedingungen werden in Bezug auf Straßen- und Lichtverhältnisse erfasst. Straßenbedingungen können verschiedene Zustände wie Überschwemmungen, Schnee, Eis, Nässe oder Trockenheit umfassen. Lichtverhältnisse werden als Tageslicht und Dunkelheit mit und ohne Beleuchtung unterschieden. Die Unfallart wird durch Parameter wie Junction Control (Vorfahrtsregelung an Kreuzungen), Junction Detail (Details zur Kreuzung), Straßentyp und mögliche Gefahren auf der Fahrbahn bestimmt. Die Schwere der Unfälle wird in Kategorien wie fatal, ernst oder leicht eingestuft und es werden Informationen über die Anzahl der Verletzten und beteiligten Fahrzeuge pro Unfall erfasst. Zusätzlich werden potenzielle Gefahren auf der Fahrbahn wie Objekte, Tiere oder vorherige Unfälle dokumentiert.

Laut Kaggle sind zu diesem Datensatz keine anderen Codes, Analysen oder Auswertungen vorhanden und stammen aus einer Sammlung von öffentlichen Unfallartikeln und Polizeiakten. Obwohl kein Autor angegeben ist, wird er Datensatz als seriös eigeschätzt.

**Data Cleaning**

Zuerst wird der Datensatz in beide Tools als Excel Datei eingelesen. Die erste Spalte "Accident\_Index" gibt hierbei eine dreizehnstellige Hexadezimalzahl als Erkennungslabel der einzelnen Unfälle an, die für uns kein Mehrwert liefert und deshalb entfernt wird. In Python erfolgt dies über drop() Funktion, während ….

Als zweiten Step korrigieren wir Rechtschreibfehler in den Spalten "Junction\_Control" und "Accident\_Serverity". Hierbei wurde sowohl "Auto traffic signal" als "Auto trafffic sigl" als auch "Fatal" als "Fetal" falsch bezeichnet. Python ersetz die Einträge der Spalten mit replace(), während in Tableau durch Anklicken der falschgeschriebenen Kategorie, diese Überschrieben werden kann. Zudem war die Werte der Spalte "Time" in Python als String dargestellt, die wir für spätere Berechnungen in das Datumsformat %H:%M:%S umgewandelt haben. In Tableau hat der Prep Builder diesen Schritt selbstständig erledigt, da er die Werte automatisch als Uhrzeit anerkannt hat. Danach erkannten Python und Tableau in der Spalte "Carriageway\_Hazards" in fast allen Zeilen angebliche Nullwerte. Diese Nullwerte stellten eigentlich dar, dass bei diesem Autounfall keine Verkehrshindernisse integriert waren und nicht das keine Werte vorhanden sind. Deshalb haben wir alle angeblichen Nullwerte der Spalte in Python mithilfe fillna() auf den Wert "None" aufgefüllt und in Tableau …

Erst daraufhin wurden alle anderen Nullwerte des gesamten Datensatzes entfernt. In Python wird dafür der Befehl dropna() verwendet währende Tableau …

Dadurch erhielten wir einen bereinigten Datensatz mit 21 Spalten und 300.000 Zeilen als CSV-Datei.

**SWOT-Analyse**

Die SWOT-Analyse stellt Stärken und Schwächen sowie Chancen und Risiken gegenüber. Dafür werden intern und extern positive und negative Aspekte gesammelt, aus denen man vier verschiedene Strategien ableiten kann. Dies vier Strategien dienen zum einen dazu, die Chancen so zu nutzen, dass die Stärken ausgenutzt und Schwächen minimiert werden. Zum anderen sollen Risiken durch Stärken vorbeugt werden und die Gefahr, dass Risiken im Bereich der Schwächen eintreten verringert wird.

Literatur: [SWOT Analyse • Beispiele, Definition, Vor- und Nachteile · [mit Video] (studyflix.de)](https://studyflix.de/wirtschaft/swot-analyse-234)

Ein Bild, das Text, Screenshot, Schrift, Logo enthält.

Automatisch generierte Beschreibung

**SWOT-Tableau**

Die Stärken von Tableau liegen bei benutzerfreundlichen Drag-and-Drop-Funktion für die Erstellung von Visualisierungen, sodass keine umfangreiche Programmierkenntnisse erforderlich sind. Darüber hinaus bietet die Plattform eine breite Palette von Visualisierungsoptionen, die es Benutzern ermöglichen, ihre Daten auf vielfältige Weise, je nach den Anforderungen und Zielen ihrer Analyse, darzustellen. Ein weiterer Vorteil der Plattform ist die Möglichkeit, viele verschiedene Datenquellen zu importieren und kombinieren. Des Weiteren enthält Tableau verschiedene Erweiterungen, wie beispielsweiße Tableau Prep, die auf die einzelnen Schritte des CRISP-DM spezialisiert sind und damit präzisere und effektivere Ergebnisse liefern.

Schwächen können in Form von Skalierbarkeitsprobleme und mangelnde Flexibilität auftreten, was den Benutzer bei der Durchführung bestimmter Analysen oder der Darstellung bestimmter Datentypen einschränkt. Ein weiteres potenzielles Hindernis ist der hohe Preis der Plattform, der für viele Unternehmen abschreckend wirkt und eine finanzielle Hürde darstellen kann.

Trotz dieser Herausforderungen bietet die Plattform auch eine Reihe von Chancen für die Zukunft. Die Integration in die Cloud bietet die Möglichkeit zur flexiblen und skalierbaren Nutzung sowie zur Speicherung und Verarbeitung großer Datenmengen. Auch der Einsatz vom KI innerhalb des Tools wird vor allem in Zukunft eine große Rolle noch spielen. Außerdem eröffnet die Integration in Big-Data-Plattformen die Möglichkeit zur Verarbeitung und Analyse von Daten aus verschiedenen Quellen und in unterschiedlichen Formaten. Darüber hinaus werden durch eine eigene Lernplattform der Einstieg und möglicher Schulungsbedarf der Benutzer gedeckt.

Trotz dieser Chancen muss die Plattform auch mit Bedrohungen umgehen, darunter die Konkurrenz durch Open-Source-Tools, die ähnliche Funktionen bieten, jedoch möglicherweise kostenlos verfügbar sind. Zudem könnten mögliche Integrationsprobleme in vorhandene Unternehmenssysteme dazu führen, dass Tableau nicht nahtlos in die bestehende IT-Infrastruktur integriert werden kann und damit die Nutzung beeinträchtigen wird, was die Attraktivität der Plattform verschlechtert.

Tableau kann sich vor allem als benutzerfreundliches und universelles Tool zu positionieren, das von Benutzern mit unterschiedlichem Kenntnisstand einfach genutzt werden kann. Durch die Betonung der Stärken und den klaren Mehrwert für die Benutzer kann die Plattform ihre Wettbewerbsposition stärken und sich von der Konkurrenz abheben. Der Preis kann dabei durch gute Leistung und vielzählige Vorteile gerechtfertigt werden und bei Privatpersonen durch die kostenlose Teilversion umgangen werden.

Ein Bild, das Text, Screenshot, Schrift, Zahl enthält.

Automatisch generierte Beschreibung

**SWOT-Python**

Python bietet mit einer großen Bandbreite an Bibliotheken, wie beispielsweiße Matplot, Seaborn oder Ploty eine unglaubliche Vielfalt an Diagrammen, um Daten auf anschauliche und aussagekräftige Weise darstellen. Diese Visualisierungen können flexibel skaliert und angepasst werden. Darüber hinaus zeichnet sich Python durch seine Effizienz bei der Verarbeitung großer Datenmengen und Umgang mit verschiedenen Datentypen aus. Zusammen mit der Fähigkeit andere Tools und Programmiersprachen zu integrieren, ist Python ein vielseitiges Werkzeug für jegliche Data-Science-Anwendungen.

Der Einstieg in Python hingegen kann für Personen ohne Programmiererfahrung eine steile Lernkurve bedeuten, was automatisch zu einem erhöhten Aufwand führt. Außerdem kann bei komplexeren Projekten die Übersichtlichkeit schnell verloren gehen.

Dennoch bietet Python zahlreiche Chancen und Möglichkeiten durch die Verfügbarkeit von Tutorials und Hilfsressourcen wie ChatGPT oder Tutorials, die es Anfängern erleichtern die Sprache zu erlernen und ihre Fähigkeiten weiterzuentwickeln. Darüber hinaus ist Python eine Open-Source-Programmiersprache, die für jeden frei zugänglich ist, was sie zu einer attraktiven Wahl für Unternehmen und Privatpersonen macht.

Einige Risiken liegen in der Komplexität, die Anfänger abschrecken könnte, sowie in der begrenzten Integration von KI-Funktionen direkt in Python.

Insgesamt ist Python ebenfalls als Universalwerkzeug anzusehen, da es vielseitig einsetzbar und ohne Kosten verbunden ist. Durch die Bibliotheken mit vielen und auch komplexen Visualisierungen können Anreize für Anfänger gesetzt werden, wobei zahlreiche Lernmittel den Lernprozess beschleunigen. Zudem muss beachtet werden, dass viele Anwender im Data Science Bereich schon Programmiererfahrungen mit sich bringen und damit die größte Schwäche von Python umgangen wird.

Ein Bild, das Text, Screenshot, Schrift enthält.

Automatisch generierte Beschreibung

**Nutzwertanalyse**

Die Nutzwertanalyse ist ein Verfahren, bei dem die Entscheidung zwischen mehreren Optionen erleichtert werden soll. Dafür müssen Kriterien und eine Skala festgelegt werden, sodass jede Option pro Kriterium ein Wert von der Skala erhält. Am Ende werden die Punkte zusammenaddiert und die Option mit der besten Punktzahl hat gewonnen.

Literatur: [Nutzwertanalyse: Erklärung und Beispiel · [mit Video] (studyflix.de)](https://studyflix.de/wirtschaft/nutzwertanalyse-315)

**Nutzwertanalyse Kategorien**

In unserem Fall erfolgt die Bewertung an einer Skala von 1 bis 5, wobei 1 die schlechteste und 5 die beste Bewertung sein kann. Das erste Kriterium ist die Benutzerfreundlichkeit, gemessen daran, ob das Tool einfach und intuitiv bedienbar ist. Die Lernkurve und Lernunterstützung bezieht sich auf die vorhanden Lernangebote und den geschätzten Zeitraum, den es zur Erlernung der benötigten Skills braucht. Je nach Anzahl der unterstützen Datenquelle und maximalen Datengröße richten sich die Kriterien Datenquellunterstützung und Datengröße. Die Integration des Tools mit anderen Plattformen beschreibt die Kompatibilität, während kollaborativ die Effektivität der Zusammenarbeit mehrerer Benutzer an einem Projekt bewertet. Für die Visualisierung gibt es vier Kategorien: Visualisierungsarten, Komplexität der Visualisierung, Anpassbarkeit und Interaktivität. Wie die Bezeichnungen der Kategorien schon verdeutlichen, wird gemessen wie viele verschiedene Visualisierungsarten es gibt, wie gut diese auch komplexe Daten Konstrukte darstellen können und ob die Diagramme anpassbar und interaktiv sind. Zuletzt wird bewertet welche Exportmöglichkeiten bestehen und wie die Preisgestaltung sich zusammensetzt.

**Nutzwertanalyse Tableau**

Tableau ist durch Drag-and-Drop sehr benutzerfreundlich mit keiner großen Lernkurve, wobei durch die eigenen Lernplattform von Tableau eine umfassende Lernunterstützung vorhanden ist. Da x verschiedene Datenquellen unterstützt werden und Daten bis x unterstützt werden, eignet sich Tableau als Plattform für (fast) alle Anwendungsfälle.

Die Zusammenarbeit von mehreren Personen am selben Projekt ….. + integration

Tableau biete mit 12 verschiedenen Visualisierungsarten eine große Bandbreite, die auch komplexe Sachverhältnisse darstellen können und zum Großteil anpassbar sind. Ebenfalls können alle Diagramme interaktiv gestaltet werden und in x verschiedene Dateiformate exportiert werden. Schlecht hervorsticht jedoch der hohe Preis mit 75 Dollar pro Person pro Monat.

**Nutzwertanalyse Python**

Bei Python stellt Benutzerfreundlichkeit und Lernkurve die größte Herausforderung dar, da das Erlernen einer Programmiersprache schwierig und lange dauert, wobei jedoch umfassende und verschiedenste Lernunterstützung geboten ist. Zudem unterstützt Python viele verschiedene Datenquellen ohne genau Obergrenze für die Größe der Datensätze. Python kann zwar kollaborativ genutzt werden durch beispielsweiße GitHub….

Python ist mit unzähligen anderen Tools und Programmiersprachen integrierbar, was eine ähnliche Kompatibilität wie bei Tableau aufweist. Zur Visualisierung besitzt Python viele verschiedene Visualisierungsarten, die auch komplex und anpassbar sind. Jedoch muss die Interaktivität sehr aufwendig implementiert werden oder ist gar nicht möglich. Auch die Exportmöglichkeiten von Python Dateien ist begrenzt. Zuletzt sticht jedoch der Preis positiv hervor, da Python für alle zugänglich und kostenlos ist.

Zusammengerechnet kommt man bei Tableau von möglichen 65 Punkte auf 54 Punkte, während Python nur 44 Punkte erreicht, Dies zeigt deutlich, dass Tableau in den meisten Aspekten gut abschneidet und deutlich besser als Python ist.

Ein Bild, das Text, Screenshot, Zahl, Schrift enthält.

Automatisch generierte Beschreibung

**Fazit**

Zusammenfassend kann gesagt werden, das Tableau sowohl in der SWOT-Analyse als auch in der Nutzwertanalyse im Vergleich mit Python sehr gut abgeschnitten hat. Data Cleaning ist durch die extra Destopanwendung Tableau Prep einfach, intuitiv, präzise und schnell, während eine Vielzahl an Visualisierungsarten um ein Vielfaches einfacher und schneller implementiert, werden können. Die Funktion der Dashboards und interaktiven Diagramme bieten somit ein hervorragendes Ergebnis unterschiedlichster Datenvisualisierung als Entscheidungsgrundlage datengetriebener Entscheidungen. Komplexe Analyse von Trends ist nicht nur für Experten, sondern auch für Anfänger keine Herausforderung. Gerade für Privatpersonen schneidet Tableau durch die kostenlose Version am besten von allen in dieser Vorlesung betrachtenden Visualisierungstools inklusive Python ab. Im Unternehmenskontext ist Tableau ebenfalls empfehlenswert, wenn man bereit ist, den hohen, aber angemessenen Preis zu zahlen und Tableau extra in die Unternehmen IT-Infrastruktur implementieret werden kann.