Unterrichtsreihe „Datenexploration - Datendetektive bei der Arbeit“

In dieser Unterrichtseinheit geht es um Data Science Inhalte für die 9./10. Klasse. Anhand eines fiktiven Rahmenbeispiels einer Online-Plattform, die Werbung für Jugendliche passgenau schalten möchte, sollen die Schülerinnen und Schüler motiviert werden, als Datendetektive in einem Datensatz nach Spuren und Mustern zu suchen, um so die Online-Plattform zu beraten.

Es stehen Arbeitsblätter, PowerPoint-Präsentationen, Anleitungen, der Datensatz (JIM-Datensatz), Variablenlisten und dieser Überblick über die einzelnen Stunden zur Verfügung. Für die Datenanalyse wird die Software CODAP ([codap.concord.org](https://codap.concord.org/app/static/dg/de/cert/index.html)) genutzt, die im Internet frei zugänglich ist.

Gearbeitet wird mit Daten von fast 500 Schülerinnen und Schülern, die zu vielen Merkmalen im Freizeit- und Medienbereich Angaben gemacht haben (JIM für Jugend-Information-Medien). Der Datensatz liegt für diese Unterrichtsreihe in zwei Versionen vor. Zum einen steht eine reduzierte Version zur Verfügung mit 50 Variablen. Zum anderen kann die Vollversion mit über 160 Variablen im Unterricht eingesetzt werden. Es wird empfohlen, aufgrund der besseren Übersichtlichkeit den Datensatz mit der reduzierten Variablenliste zu verwenden. Hier sind multivariate und interessante Entdeckungen möglich! Die Lehrkraft hat jedoch die Möglichkeit je nach Leistungsfähigkeit der Klasse/des Kurses zu differenzieren oder für besonders leistungsfähige Schülerinnen und Schüler ein binnendifferenziertes Angebot zu unterbreiten (siehe unten).

Der Link zum Datensatz, mit dem im Projekt gearbeitet wird (50 Variablen):

<https://tinyurl.com/datendetektive-50>

Die Unterrichtsreihe umfasst 8 Unterrichtsstunden. Das Herzstück der Unterrichtsreihe ist eine eigenständige Data Science Projektarbeit der Schülerinnen und Schüler in den Stunden 5 und 6 mit Präsentationen in Stunde 7. Die Projektarbeit findet in Kleingruppen statt, in denen sich die Schülerinnen und Schüler eigenständig als Data Science Experten mit dem Datensatz arbeiten. Dafür ordnen sie sich vier inhaltlichen Gruppen der JIM-Daten zu analog zu den Kundenwünschen der Online-Plattform (diese Zuordnung geschieht in Stunde 4):

* Kunde 1 möchte TikTok bewerben,
* Kunde 2 möchte LetsPlay\_YouTube Videos bewerben,
* Kunde 3 möchte Online Zeitungen bewerben,
* Kunde 4 möchte feste Spielekonsolen bewerben.

Präsentationen aller Gruppen beschließen die Gruppenarbeit. In der letzten Stunde findet eine Reflektion über das Vorgehen für die Datenexploration statt und persönliche und gesellschaftliche Aspekte können diskutiert werden.

Überblick über die Unterrichtsreihe

|  |  |  |
| --- | --- | --- |
| **Stunde** | **Thema** | **Material** |
| 1 | **Einführung**  In dieser Stunde wird in das Projekt “Datendetektive bei der Arbeit” eingeführt. Dazu wird der komplette Rahmen der Unterrichtsreihe aufgezeigt. In zwei Arbeitsphasen werden die SuS an die vorliegenden Umfragedaten und an die Datenanalyse in CODAP herangeführt.  Das Dokument Stunde\_1\_Überblick bietet Gestaltungshinweise für diese Stunde. | Stunde\_1\_Überblick  Arbeitsblatt1\_Datensatz  Variablenliste\_JIM\_2021\_reduziert |
| 2-3 | In diesen beiden Stunden geht es darum, dass die Schülerinnen und Schüler zu Data Science Experten werden, um anschließend selbständig arbeiten zu können.  **Einführung in statistische Grundbegriffe und Erwartungshaltung aufbauen**  Die SuS erhalten Informationen über grundlegende statistische Fachbegriffe (Infoblatt\_Glossar). Die kennengelernten statistischen Grundbegriffe übertragen die SuS anhand von Arbeitsblatt2\_Begriffe auf den vorliegenden Datensatz.  **CODAP und Analysemethoden**  Es wird in die Bedienung von CODAP eingeführt und Auswertungsmöglichkeiten bezüglich verschiedener Prozente werden nacheinander erarbeitet. Je nach Kurs/Klasse können sich die SuS die Auswertungsmöglichkeiten anhand der Anleitungen selbst erarbeiten und bei den Arbeitsblättern anwenden. Alternativ kann die Lehrkraft mithilfe der PowerPoint in die Analysemethoden einführen und die SuS bearbeiten dann die Arbeitsblätter. Dies ist die notwendige Voraussetzung für die eigenständige Schülerexploration in den folgenden Stunden.  Die Anleitungen für die Selbstlernphase der SuS beinhalten das Auswerten von eindimensionalen Verteilungen (Anleitung\_CODAP\_1dim) und zweidimensionalen Verteilungen (Anleitung\_CODAP\_2dim\_A und \_B).  Wir schlagen vor, mit der Anleitung (Anleitung\_CODAP\_2dim\_red) in die Exploration von zweidimensionalen Verteilungen zu starten. Hier wird zunächst von einer Variablen nur eine Ausprägung betrachtet und somit die Komplexität der Auswertung von einer 7x7 Matrix auf eine 7x1 Matrix reduziert.  **Technik**  Hier sollte den SuS auch gezeigt werden, wie sie Graphiken aus CODAP in eine Word-Datei oder eine PowerPoint Präsentation kopieren können. Ggf. ist auch das Weitergeben des CODAP-Dokuments per Link eine gute Möglichkeit, um Schülerbearbeitungen zu dokumentieren, bzw. zu kontrollieren. | Infoblatt\_Glossar  Arbeitsblatt2\_Begriffe  PowerPoint-Präsentation (optional)  Anleitung\_CODAP\_1dim  Anleitung\_CODAP\_2dim\_red  Anleitung\_CODAP\_2dim\_A  Anleitung\_CODAP\_2dim\_B  Arbeitsblätter 3 – 5  Anleitung\_Graphikenkopieren  Oder YouTube-Video zum Kopieren des Arbeitsstands per Link:  <https://youtu.be/otLuX8hhtq8> |
| 4 | **Erwartungshaltung aufbauen und passende Fragen stellen**  In dieser Stunde werden die Schülerinnen und Schüler in Kleingruppen für die komplette restliche Unterrichtseinheit eingeteilt und vier verschiedenen inhaltlichen Bereichen der Daten/Kunden der Online-Plattform zugeordnet (Arbeitsblatt3\_Gruppenbildung). Anhand von AB3 findet eine theoretische Auseinandersetzung mit den Daten, resp. der Variablenliste, statt, um eine Erwartungshaltung für die kommenden Stunden aufzubauen. Diese Erwartungshaltung ist zentral für die eigene Datenanalyse und soll auf Plakaten festgehalten werden. Die hierzu erstellten Plakate sollten im Verlauf der weiteren Unterrichtseinheit sichtbar für alle SuS im Klassenraum hängen.  Die Erfahrung hat gezeigt, dass das Stellen von geeigneten (statistischen) Fragestellungen für SuS eine Hürde darstellt, deshalb liegt ein weiterer Fokus dieser Stunde darauf, die SuS in Kleingruppen geeignete Fragen erarbeiten zu lassen. Hierfür wird die Methode Think-Pair-Share genutzt. | Arbeitsblatt6\_Gruppenbildung  Plakate  PowerPoint als Hintergrundinformation für L  Arbeitsblatt6a\_Fragen |
| 5-6 | **Projektarbeit in Kleingruppen**  In diesen beiden Stunden wird das Vorgehen der Datenexploration von den Kleingruppen zunächst geplant (Arbeitsblatt 7). Anschließend findet die Datenexploration in den Kleingruppen statt (Arbeitsblatt 8) und die Präsentation der Ergebnisse wird vorbereitet. Zu Beginn der sechsten Stunde können evtl. Kriterien für gute statistische Präsentationen besprochen werden, um die Erstellung der Präsentationen vorzubereiten (Leere Power-Point-Vorlage). Die Lehrkraft steht in den Stunden 5-6 vor allem unterstützend für Fragen und Anregungen zur Verfügung, die SuS arbeiten möglichst selbständig mit CODAP und PowerPoint. | Arbeitsblätter 7-8  Anleitung\_HinweisePräsentation  Leere PowerPoint-Vorlage für Schülerpräsentationen |
| 7 | **Präsentationen der Kleingruppen**  In dieser Stunde finden die Präsentationen der Schülerergebnisse statt. Dabei kann immer eine Gruppe einen speziellen Feedback Auftrag bekommen, damit eine inhaltliche Diskussion angeregt wird (Arbeitsblatt 9). Anhand von Arbeitsblatt 10, das sich gut als Hausaufgabe und Ergebnissicherung eignet, können die SuS überprüfen, ob sie eine Datenanalyse durchführen können. | PowerPoint-Präsentationen der SuS  Arbeitsblatt 9-10 |
| 8 | **Reflektieren**  In dieser Stunde findet eine gemeinsame Reflexion des gesamten Projekts statt. Die einzelnen Schritte der durchgeführten Datenanalyse werden den Stationen des PPDAC-Kreislaufs zugeordnet (Powerpoint\_Präsentation).  Weiter können unter anderem persönliche und gesellschaftliche Auswirkungen einer Datenexploration angesprochen werden und versucht werden, „über den Tellerrand” zu schauen. | Powerpoint\_Präsentation |

**Weitere Hinweise**

## Möglichkeit zur Differenzierung:

Die Unterrichtsreihe ist so angelegt, dass mit einem „kleinen“ Datensatz gearbeitet werden kann. Dieser enthält 50 Variable und bietet reichhaltig Möglichkeiten für Explorationen.

Für besonders motivierte Schülerinnen und Schüler kann stattdessen auch mit dem großen Datensatz gearbeitet werden, der alle 160 Variablen erhält, die in der Umfrage erhoben wurden. Dies erfordert jedoch ein hohes Engagement und gutes Arbeiten mit der Variablenliste seitens der Schülerinnen und Schüler!

## Zum Umgang mit den Aufgaben:

In der Unterrichtsreihe wird viel mit Arbeitsblättern gearbeitet. Um den Lernprozess zu dokumentieren und gleichzeitig die Motivation hoch zu halten, kann man die Aufgaben auch direkt in einer PowerPoint-Präsentation erarbeiten lassen. Neue Aufgaben können auf neuen Folien bearbeitet werden und am Ende können aus den verschiedenen Schulstunden Ergebnisse genommen werden, um die Abschlusspräsentation zu gestalten.

## Zum Umgang mit CODAP:

CODAP ist in weiten Teilen auf Deutsch verfügbar. Sollte die Sprache nicht auf Deutsch eingestellt sein, so kann rechts oben die Sprache umgestellt werden.

Link zum Datensatz mit allen 161 Variablen:

<https://tinyurl.com/datendetektive>