**2. Meet SQL**

* EMPFHOLENE KURSE & ONLINE ENVIRONMENTS

Nutze zum Einstig in SQL (einen der) folgende(n) E-Learning Kurse:

* + [**SQL Tutorial | w3schools**](https://www.w3schools.com/sql/)
  + [**ER Diagrams | coursera**](https://www.coursera.org/lecture/analytics-mysql/how-entity-relationship-diagrams-work-6kP20)
  + [**Learn SQL | codeacademy**](https://www.codecademy.com/learn/learn-sql)
  + [**Introduction to Databases and SQL Querying | udemy**](https://www.udemy.com/course/introduction-to-databases-and-sql-querying/?LSNPUBID=JVFxdTr9V80&ranEAID=JVFxdTr9V80&ranMID=39197&ranSiteID=JVFxdTr9V80-jFWXSL04J9C0aInGIEOp7A&utm_medium=udemyads&utm_source=aff-campaign)

THEORETHISCHE FRAGESETLLUNGEN

* + Was ist eine relationale Datenbank?
  + Was sind die Vorteile von Datenbanken im Vergleich zu FileSystems?
  + Was versteht man, wenn man über Redundanz von Daten spricht?
  + Was ist und wie macht man einen ER-Diagram?
  + Was sind die Entitäten, Attributen und Beziehungen?
  + Welche Arten von Beziehungen gibt es?
  + Was ist ein DBMS?
  + Was sind die gängigste SQL Management CLI Befehle?
  + Wie erstellt man die Benutzer und vergibt die Rechte?
  + Wie ist eine Datenbank organisiert?
  + Was ist die Relation in einer Datenbank?
  + Was sind die Tuples?
  + Was sind die Attributes?
  + Was ist ein Primärschlüssel in Bezug auf Datenbanken?
  + Wie wählt man einen Key?
  + Was sind SQL Queries?
  + Was sind gängige SQL Query Commands?
  + Was ist JOIN Statements und welche Typen von JOINS gibt es?

VERPFLICHTENDES PRAXISBEISPIEL

Zur Übung der SQL Basics setze dich mit der folgenden Aufgabenstellung auseinander und führe die einzelnen Schritte aus. Dokumentiere dabei dein Vorgehen.

**Vorbereitung:**

* + - Installiere einen lokalen PostgreSQL Server.
    - Stelle die Erreichbarkeit sicher durch einen Test-Login.
    - Erstelle einen neuen Benutzer und gib ihm die Rechte für die Erstellung und Verwaltung einer Datenbank.
    - Verwende nur den neu erstellten Benutzer für die restliche Aufgaben.

**Szenario und Data:**

Unsere Firma hat einen wichtigen Kunde in Boston, MA, USA. Die Firma heißt AlphaTech Consulting und ist eine der bekanntesten Tech Consulting Firmen in den USA - besonders unter den Silicon Valley Unternehmen. Unsere Firma gilt dort als Experte für Daten und Datensicherheit.  
AlphaTech hat seit kurzem eine neue Managerin, die sich gerne einen Überblick über die Mitarbeiter:innen, Departments, Kunden und Projekte verschaffen möchte. Da ihre Mitarbeiter\*innen aktuell sehr beschäftigt mit ihren laufenden Projekten sind, hat sie uns ersucht, die, in dem anbei vorgelegtem .csv File liegende Daten in einer Datenbank zu organisieren und danach ein paar Insights für sie zu erwerben. Unser Management ist zuversichtlich, dass du mit deinen frisch erworbenen SQL Kenntnissen die richtige Person für diese Aufgabe bist.  
Wir zählen auf dich!

Daten:[**AlphaTechConsultingEmployees.csv**](https://learnhub.bearingpoint.com/pluginfile.php/3199/mod_unilabel/intro/AlphaTechConsultigEmployees.csv?time=1682351862998)

**1 ) Vorbereitung und Erstellen eines ERD - Entity Relationship Diagram**

* + - * Schaue dir die Daten an. Welche Entities, Attributes und Relations gibt es? Welche Types von Relations sind es?
      * Erstelle ein ERD mittels [**draw.io**](https://app.diagrams.net/).
      * Bedenke beim Erstellen des ERD die Vorgaben am Anfang des Beispiels und welche Daten einzigartig sind und auch welche Daten miteinander verknüpft werden sollen.
      * Exportiere dein ERD als .png file und füge die Datei zu deiner Dokumentation in Confluence hinzu.

**2) Eine Datenbank mithilfe des ERD erstellen**

* + - * Aufgrund des ERDs: Wie viele Tabellen brauchst du in der Datenbank?
      * Welche Attribute gehören zur welcher Tabelle?
      * Wie sind die Tabellen mit einander zu verknüpfen?
      * Erstelle ein SQL Script (oder auch ein Python Script um das ganze zu automatisieren) zur Erstellung der Datenbank.
      * Führe dein Script aus und erstelle die Datenbankstruktur.
      * Überprüfe mittels psql CLI ob die Datenbank mit allen Tabellen und entsprechenden Attributen wirklich vorhanden ist.

**3) Fülle die Datenbank mit den Daten aus**

* + - * Schreibe eine Python Script, welches die Daten aus dem csv File einliest, die Daten sortiert, sich mit der Datenbank verbindet und die Daten in den entsprechenden Tabellen speichert.
      * Führe das Script aus und fülle die Datenbank mit den Daten aus.
      * Überprüfe mittels psql CLI ob die Daten wirklich in der Datenbank sind.

**4) Erwerbe die Insights anhand SQL Queries**Wie schon erwähnt, sollen wir für den Kunden ein paar Insights über dessen Mitarbeiterstruktur und Projekte liefern.

* + - * Überlege dir ein Design von SQL Queries die Antworten auf die untenstehenden Fragen 'entdecken' können. (eine manuelle Beantwortung ist vom Kunden nicht gewünscht)
      * Erstelle ein Python Script das SQL Queries ausführt und die Ergebnisse in einer Datei auf anschauliche und übersichtliche Art und Weise speichert.
      * Führe das Script aus und lade das Ergebnis zusammen mit deiner Dokumentation in Confluence hoch.

Welche Mitarbeiter:innen arbeiten im Sales? Wie heißen sie, wann sind sie geboren und welches Geschlecht haben sie?

Übersicht zur Gleichstellung der Mitarbeiter:innen: Wie groß ist der Anteil der weiblichen Kolleginnen? Verdienen diese im Durchschnitt gleich viel wie ihre männlichen Kollegen?

Wie viele Mitarbeiter:innen sind in den letzten 3 Jahren neu angestellt worden?

Wie alt ist der/die älteste/r Mitarbeitende? Was ist das Durchschnittsalter in der Firma?

Wie hoch ist die Gehaltssumm der Mitarbeiter\*innen die für Projekte von Apple arbeiten?

Welche Mitarbeiter:innen arbeiten an den Projekten 'Google Cloud Platform' und 'AWS'? Wie heißen sie, welche Position haben sie und in welchem Department sind sie tätig?

Welche Positionen gibt es in dem 'Software Engineering' Department?

Wie viele Manager gibt es? Wie heißen Sie? Welche Positionen haben sie und was verdienen sie?

ANMERKUNGEN  
👉 Deine Dokumentation soll eine prägnante:

* + - * Beschreibung deiner Vorgehensweise (welche Probleme hast du definiert? welches Problem hast du auf welche Art und Weise gelöst?)
      * Begründung deiner wichtigsten Entscheidungen (Welche Alternativen hattest du und wieso hast du dich für genau die entschieden) sein
      * How-To-Anleitung (für deine Kolleg:innen wenn sie deine Scripts verwenden wollen) darstellen.