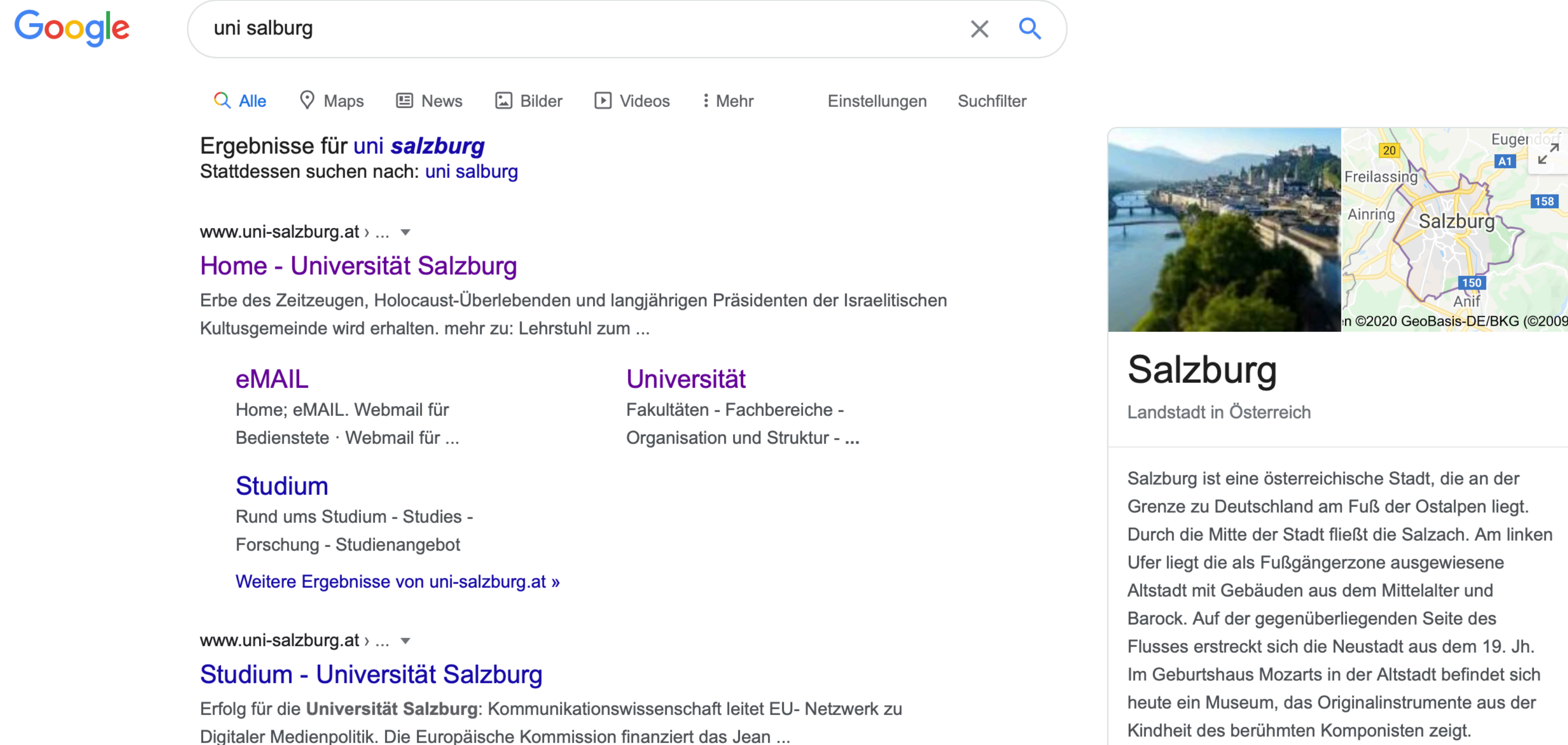


PS Ähnlichkeitssuche in großen Datenbanken

Introduction

Motivation

Introduction



The image shows a Google search interface. The search bar contains the text "uni salzburg". Below the search bar, there are navigation options: "Alle", "Maps", "News", "Bilder", "Videos", "Mehr", "Einstellungen", and "Suchfilter". The search results are for "uni salzburg". The first result is from "www.uni-salzburg.at" with the title "Home - Universität Salzburg". The snippet for this result reads: "Erbe des Zeitzeugen, Holocaust-Überlebenden und langjährigen Präsidenten der Israelitischen Kultusgemeinde wird erhalten. mehr zu: Lehrstuhl zum ...". Below this result, there are two columns of information: "eMAIL" (Home; eMAIL. Webmail für Bedienstete · Webmail für ...) and "Universität" (Fakultäten - Fachbereiche - Organisation und Struktur - ...). Below these, there is a "Studium" section (Rund ums Studium - Studies - Forschung - Studienangebot) and a link "Weitere Ergebnisse von uni-salzburg.at »". A second result is also from "www.uni-salzburg.at" with the title "Studium - Universität Salzburg" and a snippet: "Erfolg für die Universität Salzburg: Kommunikationswissenschaft leitet EU- Netzwerk zu Digitaler Medienpolitik. Die Europäische Kommission finanziert das Jean ...". To the right of the search results is a knowledge panel for "Salzburg". It includes a photograph of the city, a map, and a text description: "Salzburg ist eine österreichische Stadt, die an der Grenze zu Deutschland am Fuß der Ostalpen liegt. Durch die Mitte der Stadt fließt die Salzach. Am linken Ufer liegt die als Fußgängerzone ausgewiesene Altstadt mit Gebäuden aus dem Mittelalter und Barock. Auf der gegenüberliegenden Seite des Flusses erstreckt sich die Neustadt aus dem 19. Jh. Im Geburtshaus Mozarts in der Altstadt befindet sich heute ein Museum, das Originalinstrumente aus der Kindheit des berühmten Komponisten zeigt."

Source: <https://www.google.com/>

Motivation

Introduction

Strings

Uni Salzburg

vs.

Uni Salburg

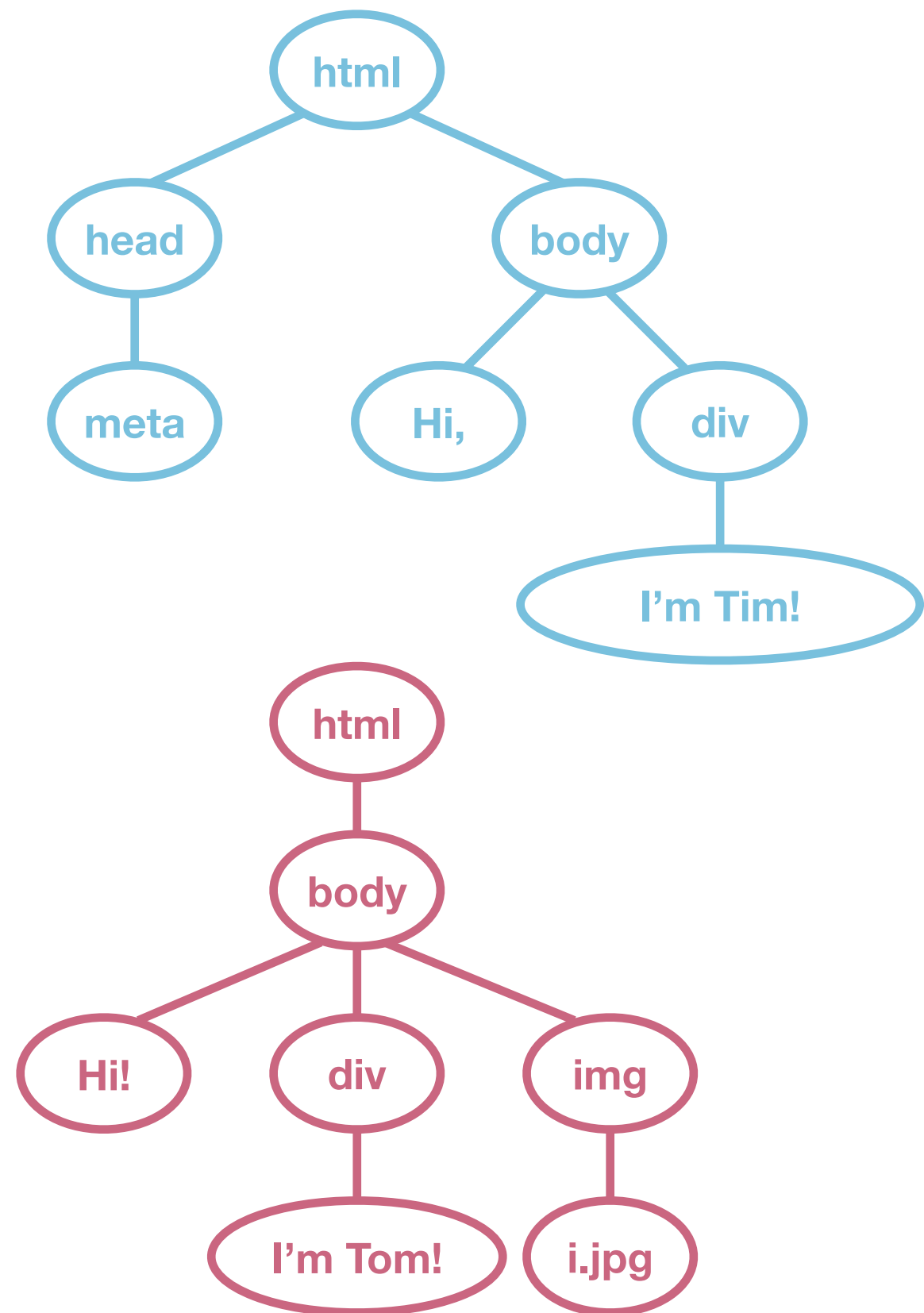
Sets

{Dog, Cat, Mouse}

vs.

{Cat, Fish, Bird}

Trees



Timetable

Introduction

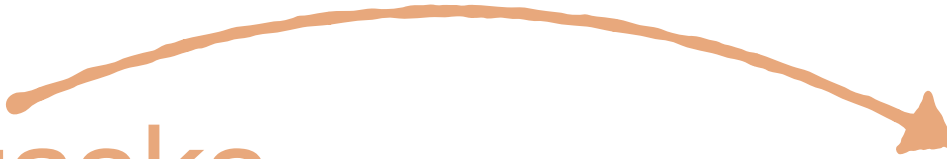
Date	Topic
01.10.2020	Introduction
08.10.2020	Set Similarity Joins
15.10.2020	Task 1
22.10.2020	Task 1 On-Demand Session
29.10.2020	Task 1 On-Demand Session
05.11.2020	Task 2
12.11.2020	Task 2 On-Demand Session
19.11.2020	Task 2 On-Demand Session
26.11.2020	Task 3
03.12.2020	Task 3 On-Demand Session
10.12.2020	Task 3 On-Demand Session
17.12.2020	Task 4
07.01.2021	Task 4 On-Demand Session
14.01.2021	Task 4 On-Demand Session
21.01.2021	Final Session
28.01.2021	— — —

- **Thursdays, 17:00 - 18:00.**
- **Meeting link: see e-mail.**

Mode

Introduction

- Work in groups of **up to 3** students.
- Each group has to work out **four tasks**.
- All tasks result in an executable program that has to be submitted via <https://abgaben.cosy.sbg.ac.at>.
- A detailed list of submission requirements can be found at the course website <https://dbresearch.uni-salzburg.at/teaching/2020ws/ssdb/>.



Task	Deadline
Task 1	03.11.2020, 23:55
Task 2	24.11.2020, 23:55
Task 3	15.12.2020, 23:55
Task 4	19.01.2021, 23:55

Grading

Introduction

- For each of the four tasks, the maximum number of points is 3.
 - Correctness: 1.5 points
 - Overall code quality: 1.5 points (e.g., data structures, comments)
- This results in the following grading scheme:

Points	Grade
≥ 10.5	1
(9, 10.5]	2
(7.5, 9]	3
(6, 7.5]	4
< 6	5

Communication

Introduction

- Register to the Slack and post your questions in channel [#ssdb-ps-2020ws](#).
- Attend the optional on-demand sessions.
- In important or urgent cases, write me an E-Mail.

Homework

Introduction

- Form groups with a maximum of 3 students.
- Send me one e-mail per group that lists
 - all members of the group (names and IDs) and
 - the programming language you plan to use.
- Deadline: 07.10.2020, 23:55