

Universität Bielefeld

(Fakultät für Wirtschaftswissenschaften)

Abschlussarbeit 1

(im Studiengang Wirtschaftswissenschaften)

Modul: Computergestützte Methoden

vorgelegt von *Kevin Wilm Schlieve und Timon Gerner*

Matrikel-Nr.: *4249212 und 2886213*

Geprüft durch *Prof. Dr. Dietmar Bauer und Dr. Meik Teßmer*

Bielefeld, im Dezember 2023

Contents

Contents	2
1 Bearbeitung der Aufgaben	3
1 Einleitung	3
2 Beschreibung der Daten	3
3 Berechnung der höchsten mittleren Temperatur in Excel	3
4 Datenhaltung	4
5 Berechnung der höchsten mittleren Temperatur in SQLite	4
Quellenverzeichnis	5

1

Bearbeitung der Aufgaben

1 Einleitung

Im Rahmen der Veranstaltung computergestützte Methoden haben wir uns mit den Thematiken Datenverarbeitung und Datenhaltung auseinandergesetzt. Als Anleitung um uns mit den Themen besser vertraut zu machen, haben wir unter anderem aufgetragen bekommen, die höchste mittlere Temperatur aus dem Capital Bikeshare Datensatz *1 in Verbindung mit Wetterdaten *2, zu bestimmen, einmal als Tabellenkalkulation und einmal als SQL-Abfrage. In dem folgenden LaTeX-Dokument *3 werden die Ergebnisse unserer Lösung, welche aus 2 Excel-Tabellen, einer Umletdatei und einer Pythondatei besteht, vorgestellt. Wir haben alles bei Github *4 hochgeladen, um dies vollumfänglich zugänglich zu machen. Im Folgenden beschränken wir uns deshalb auf die wichtigsten Ergebnisse und deren Erklärung.

2 Beschreibung der Daten

Die Daten lagen in Form einer csv-Datei vor, somit also kaum lesbar, dies haben wir festgestellt, als wir uns den Datensatz im Editor angesehen haben. Daraufhin haben wir den Datensatz in Excel importiert und nach Spalten aufgetrennt.

Es fällt sofort auf, dass alle Spaltennamen auf englisch geschrieben sind und Dezimalzahlen dementsprechend mit Punkten statt Kommata getrennt wurden. Bei späteren Auswertungen muss hier also zum Beispiel von Grad Fahrenheit zu Grad Celsius umgerechnet werden.

3 Berechnung der höchsten mittleren Temperatur in Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	date	station	count	wind_speed	precipitation	snowfall	snow_depth	mean_temperature	max_temperature	min_temperature	höchste mittlere Temperatur in Grad Celsius		
314	2022-07-23	18th St & Wy	95	7.38	0	0	0	86	96	77	30		
408	2022-08-09	18th St & Wy	132	7.61	0	0	0	86	97	77	30		
507	2022-06-01	18th St & Wy	68	6.93	0	0	0	85	94	74	29,4444444		

Fig. 1.1 Berechnung der höchsten mittleren Temperatur in Excel

Hierfür haben wir die Spalte "station" auf "18th St & Wyoming Ave NW" beschränkt und die Spalte "mean_temperature" in absteigender Reihenfolge sortiert. Das Ergebnis, 86 Grad Fahrenheit wurde dann per "`=UMWANDELN(86;"F";"C")`" *5 in Excel umgerechnet. Die Sortierung zeigt dann, dass es 3 Zeilen ohne "mean_temperature" Werte gibt. Diese haben wir ignoriert.

Die höchste mittlere Temperatur für "18th St & Wyoming Ave NW" betrug am 23.07.2022 und am 09.08.2022 jeweils 30 Grad Celsius.

4 Datenhaltung

Für die Erstellung unseres Datenbankschemas haben wir die Richtlinien der 1. und 2. Normalform genutzt. Die 1. Normalform *6 besagt, dass alle Informationen atomar vorliegen müssen, die 2. Normalform *7 besagt, dass zwischen den Tabellen Primär- und Fremdschlüsselbeziehungen *8 vorliegen müssen.

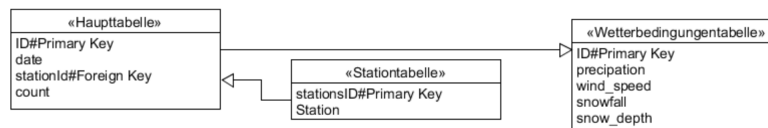


Fig. 1.2 Datenbankschema in 2.Normalform

Weiterhin haben wir uns dazu entschieden, Stationen per Zahlen zu referenzieren. So können wir zukünftige zusätzliche Informationen in Abhängigkeit der Zahl speichern und die Haupttabelle unberührt lassen.

5 Berechnung der höchsten mittleren Temperatur in SQLite

Die in Figur 1.2 gezeigten Tabellen haben wir in Python per SQL Einbindung erstellt und befüllt.

Durch die zuvorige Trennung in themenbezogene Spalten, müssen wir nun die Wetterbedingentabelle mit der Haupttabelle verbinden, dies wurde per "JOIN"- Befehl gelöst. Weiterhin schließen wir leere Zeilen aus, indem wir "NA"-Werte verbieten und filtern nach 35, unsere Zahl für unsere gruppenspezifische Station.

```

47 # Befehl für A4, hier wird max aus mean_temp gesucht, mit unserer Station (35 nach Update) außerdem wird NA als Wert ausgeschlossen
48 sql_info_max_mean_temp = """
49 SELECT MAX(mean_temperature) AS hoechsteMittlereTemperatur, main_table.station, main_table.id, weatherconditions_table.mean_temperature,
50 weatherconditions_table.wid, station_table.SID, station_table.Bezeichnung
51 FROM main_table
52 JOIN station_table ON main_table.station = station_table.SID JOIN weatherconditions_table ON main_table.id = weatherconditions_table.wid
53 WHERE mean_temperature <> 'NA' AND main_table.station = 35;
54 """

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

Die höchste mittlere Temperatur beträgt 30.0 Grad Celsius.

Fig. 1.3 Ergebnis des SQLite Befehls in Python in VSCode

SQL ▾						
< 1 / 1 > 1 - 1 of 1						
hoechsteMittlereTemperatur	station	id	mean_temperature	wid	SID	Bezeichnung
86	35	20118	86	20118	35	18th St & Wyoming Ave NW

Fig. 1.4 Ergebnis des SQLite Befehls in Python in VSCode

Quellenverzeichnis

"=UMWANDELN(86;"F";"C")", , \bibnamedelimhttps://support.microsoft.com/de-de/office/umwandeln-funktion-d785bef1-808e-4aac-bdcd-666c810f9af2 zuletzt besucht am 07.11.2023 (*5). *Microsoft support.*

1. Normalform, <https://www.datenbanken-verstehen.de/datenmodellierung/normalisierung/erste-normalform/> zuletzt besucht am 03.12.2023 (*6). *Datenbanken verstehen.*

2. Normalform, <https://www.datenbanken-verstehen.de/datenmodellierung/normalisierung/zweite-normalform/> zuletzt besucht am 03.12.2023 (*7). *Datenbanken verstehen.*

Capital Bikeshare Datensatz, <https://capitalbikeshare.com/system-data> zuletzt besucht am 10.11.2023 (*1). *Capital Bikeshare Datensatz.*

Github, https://github.com/TGerner/CoMet_Abgabe1_1 zuletzt besucht am 04.12.2023 (*4). *GitHub Repository.*

LaTeX-Dokument, LaTeX-Template von Chloé Goupy <https://www.overleaf.com/latex/templates/thesis-layout-upariscite/qvmbkpnxvtdn> zuletzt besucht am 30.11.2023

Quellen in LaTeX einfügen <https://libguides.rhul.ac.uk/referencing/latex> zuletzt besucht am 03.12.2023

Links in LaTeX einfügen <https://www.heise.de/tipps-tricks/LaTeX-URL-einfuegen-4974595.html> zuletzt besucht am 03.12.2023

Bilder in LaTeX einfügen <https://www.heise.de/tipps-tricks/LaTeX-Bilder-einfuegen-so-geht-s-4404598.html> zuletzt besucht am 03.12.2023 (*3). *Sammlung der genutzten LaTeX-Quellen.*

Primär- und Fremdschlüsselbeziehungen, <https://cloud.google.com/spanner/docs/foreign-keys/how-to?hl=de> zuletzt besucht am 03.12.2023 (*8). *Wie funktionieren Primär- und Fremdschlüssel.*

Wetterdaten, <https://www.ncei.noaa.gov/cdo-web/search?datasetid=GHCND> zuletzt besucht am 10.11.2023 (*2). *Wetterdaten.*