

Dashboard Design- Aufgabenblatt 2

Einlesen der Daten | Diagramme und Visualisierungen mit Plotly

Git: <https://github.com/hon3ypi/DashboardDesignFS23>

Aufgabe 3

a)

```
#Aufgabe 3a
import pandas as pd
import plotly.express as px

#path
path = "C:\\Users\\wanda\\01_Dashboard_Design\\Goal5.5.2_csv.csv"

import dash
from dash import dcc, html
from dash.dependencies import Input, Output

datafile = pd.read_csv(path)

print(datafile.head)
```

b)

```
#Aufgabe 3b
import pandas as pd
import plotly.express as px

#path
path = "C:\\Users\\wanda\\01_Dashboard_Design\\Goal5.5.2_csv.csv"

import dash
from dash import dcc, html
from dash.dependencies import Input, Output

datafile = pd.read_csv(path)

swiss_data = datafile[datafile['GeoAreaName'] == 'Switzerland']
pos_data = datafile[datafile['SeriesDescription'] == 'Proportion of women in managerial positions (%)']

#Array mit Länder
countries = ['Switzerland', 'Germany', 'France', 'Austria', 'Italy']
years = [2010, 2011, 2012, 2013, 2014, 2015]
#Daten filtern
filteredData = datafile.query('GeoAreaName in @countries')
filtYear = datafile.query('TimePeriod in @years')

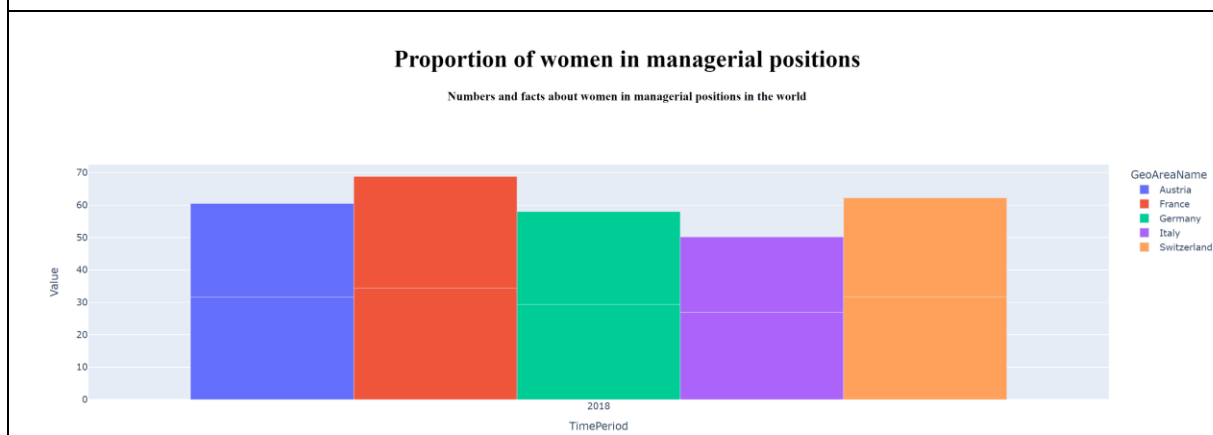
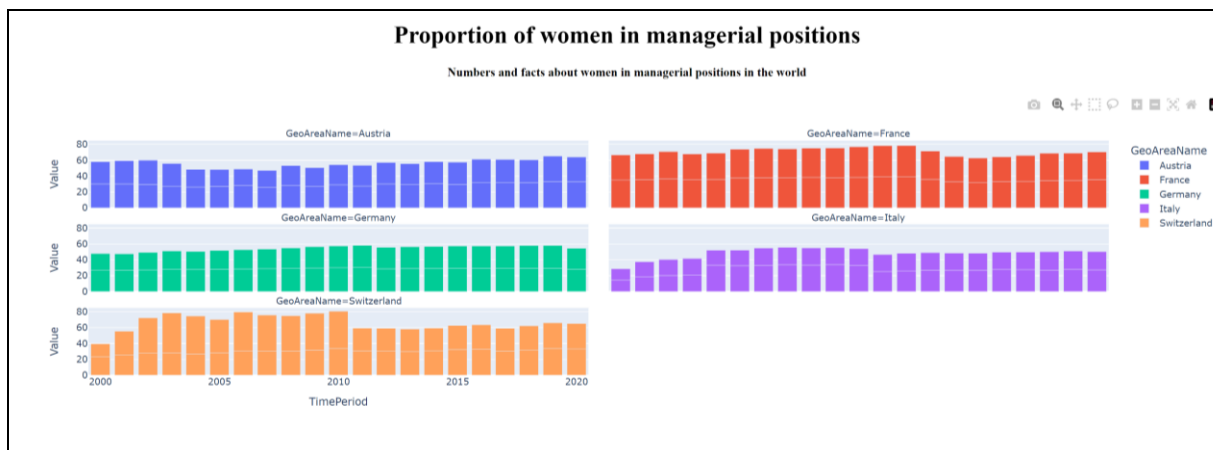
filtComb = datafile.query('GeoAreaName in @countries & TimePeriod in @years')

#print(swiss_data)
#print(filteredData)
#print(pos_data)
print(filtComb)
```

Für welche Filteroptionen haben Sie sich entschieden?

Wir haben uns Gedanken gemacht und für diese Aufgaben uns eingeschränkt. Daher haben wir uns an die Infografik von Wanda angelehnt. Die Nachbarländer der Schweiz haben wir in einem Array gespeichert und die Zeitspanne 2010 bis 2015. Natürlich könnten wir uns noch weiter einschränken oder ausbauen. Wir haben noch weitere Filteroptionen ausprobiert, aber diese sind auf dem lokalen Rechner abgelegt.

Aufgabe 4 (a und b)

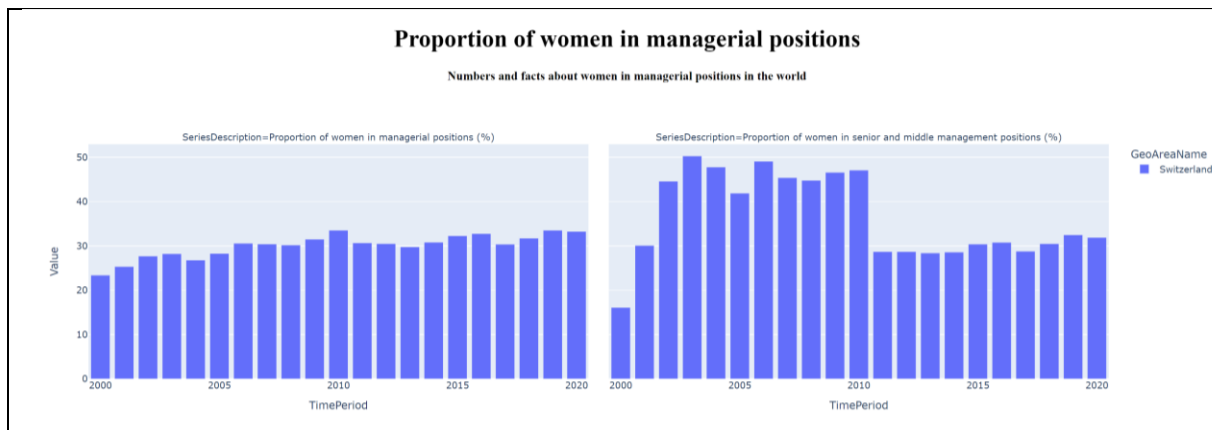


Wir haben wieder unsere Nachbarländer genommen und einen Balkendiagramm erstellt. Dabei ist uns aufgefallen, dass der Datensatz zwei Berufsbeschreibungen gibt, nämlich «Proportion of women in managerial positions (%)»Proportion of women in senior and middle management positions (%)). Dieser Datensatz müssen wir noch weiter filtern oder gar bereinigen (da einige Spalten für uns nicht relevant sind).

Beim zweiten Diagramm haben wir uns auf das Jahr 2018 eingeschränkt.

Interpretation: Es werden Daten zu 5 Länder über 20 Jahren angezeigt. Es ist eine Unterteilung vorhanden, aber die Value ist hier nicht definiert, was es ist – wir wissen jedoch, dass dies der Prozentsatz ist. So kann vermeintlich gelesen werden, dass es einen gesamten Prozentsatz ist – was demnach nicht so ist.

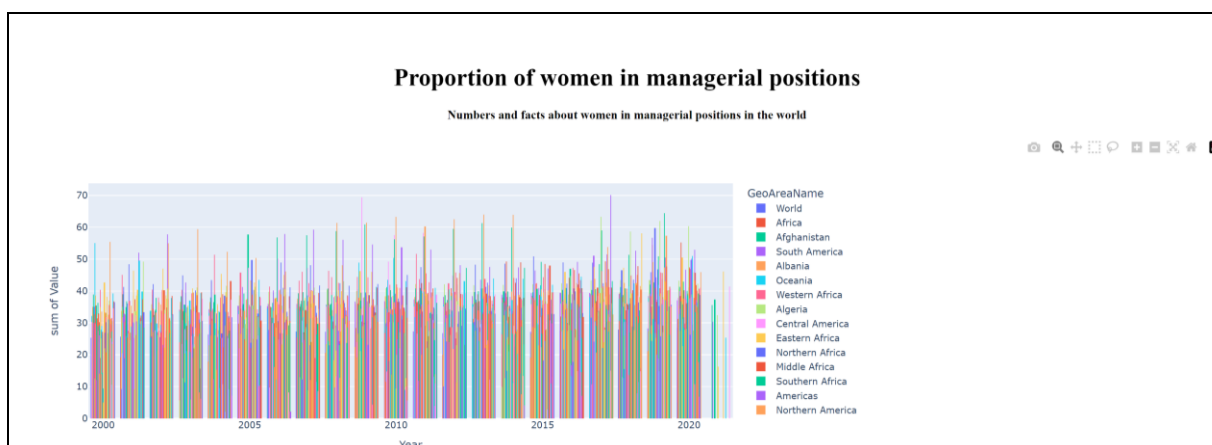
Verbesserung: Wir müssen zwingend an der Beschriftung arbeiten und die Darstellung ändern. Zudem wäre es sicher sinnvoller gewesen, in ein Diagramm zu packen als 5 einzelne zu haben – für den Vergleich.



Die zwei Berufsbeschreibungen haben wir für die Schweiz unterteilt.

Interpretation: Es sind zwei Diagramme zu sehen, die unterschiedlich sind. Auch hier bei Value ist nicht klar beschriftet, dass es die Prozentsätze sind. Der Zeitraum sind wieder 20 Jahre (2000 – 2020). Zu sehen ist ein Unterschied zwischen Frauen in Management Positionen und Frauen in höheren und mittleren Management Positionen.

Verbesserung: Prozentsatz anzeigen. Ebenfalls in ein Diagramm zusammenfassen und mit Farben arbeiten.



Hier wurde der Datensatz bereinigt und nur mit «Proportion of women in managerial positions (%)» gefüllt. Natürlich ist diese Ansicht total suboptimal.

Interpretation: Es sind zu viele Geo Area zu sehen. Es sind für jedes Jahr endlich viele Einträge, die es unmöglich es in dieser grossen Ansicht etwas genaues rauszulesen.

Verbesserung: einzelne Länder spezifisch wählen können oder das Jahr; Die Reihenfolge könnte auch angepasst werden. Eigentlich macht dieses Diagramm mässig Sinn. Ebenfalls wieder die Beschriftung und der Prozentsatz – wobei es hier keinen Sinn macht bei so viele Daten.

Disclaimer: Der Kurs ist sehr spannend, nur konnten wir nicht alle Verbesserungen machen. Zudem musste bei einzelnen die Entwicklungsumgebung aufgesetzt werden. Für die Aufgabenblätter haben wir uns bewusst eingeschränkt, da wir theoretisch endlich Lösungen für die Aufgaben hätten.