Hello and welcome to our presentation on the influence of Life Style Factors to Cardio Vascular disease.

**CLICK**

A few words about us:

We are a data analysis team working on behalf of an NGO with the aim of identifying key risk factors for CVD and present data driven strategies to reduce the risk within the US population. Therefore, we analyzed health related data to detect high-risk lifestyle patterns and provide insights for specific demographic groups.

**CLICK**

Our Data Preparation journey started with data from the CDC, specifically the national health and nutrition examination study.

**CLICK**

The first look always starts with excel or a similar table managing software. We didn’t include the sampling weights factor however, so our results might be subject to sampling bias.

**CLICK**

We cleaned the dataset and merged datasets on the Sequence number with pandas and Python.

**CLICK**

After cleanup we uploaded our data to SQL, and combined the datasets with DBEAVER of the different time periods and were able to manage it with.

**CLICK**

In the last step we conducted a RF analysis with python. All other visualizations were made with Tableau.

**CLICK**

So the first question is: What causes high blood pressure, and why is it important for cardiovascular health?

High blood pressure is the largest indicator for CVD. It can cause heart attacks which in turn are the leading cause of death both in the US and globally. Causes for high blood pressure can be corroded or obstructed arteries.

**Click**

So, from a data point of View, some problems are too complex to look at features individually. Such a problem can become easier to manage with a ML approach such as random forest.

Random Forests are a supervised machine learning algorithm for classification. It’s based on binary classifiers, aka. Decision trees. The way these trees are trained, the features with the most impact to the outcome show higher feature importance.

We yielded a classification accuracy of above 86 % for both training and test sets. Used parameters can be seen on the left. For you to take away: we assigned the outcome “hypertonicBP” to individuals with a higher than 85 mm of mercury diastolic BP and higher than 130 mm of mercury systolic BP because these are the values for hypertension in the scientific world. Here we can see the measures for optimal, normal and hypertensive BP ranges. BP is measured as systolic over diastolic BP. The optimal range poses no risk of CVD. The “normal”-declared range is the most common, although it is associated with very slight risk of causing CVD. Finally, there is hypertension, which we used our classification on.

Our analysis yielded the top 5 indicators the factors for hypertension: Age, Weight related markers such as waist circumference, weight, BMi and physical activity, shown here as the sum of days where physical activity was performed by an individual.

Click Now Kenan will talk about the influence of Age to BP.

Kenan 6 folien **CLICK** by “Previously”; **CLICK** “what does it mean to be overweight”

Harun 2 Folien

Soeren: rauchen & drinking

On this dashboard BP as well as heart rate to specific metrics of smoking and drinking are shown. We identified that smoking shows a significant increase in systolic bp over the number of cigarettes smoked per day. Interestingly both the average systolic and diastolic bp seems to fluctuate more with 8 or more cigarettes. Pulse however doesn’t seem to be as influenced by smoking. On the right sight we observed that more than 5 alcoholic beverages per week significantly increase the systolic BP.

**Click**

Let us summarize.

RF reached an 86 % classification accuracy and 15 potential predictors.

These were separated into weight related factors such as BMI, weight and waist circumference, Life style factors such physical activity during the day, Certain blood cholesterol and diet related factors such as total carbohydrate, fat and protein-intake,

Our EDA results showed that BMI and waist circumference have significant influence on blood pressure across all ethnicities. Women over 60 show higher BP than men. Our assumption is, that due to menopause, the protective power of estrogen against hypersensitivity is reduced.

Also, we find that men’s blood pressure increases more rapidly compared to women in younger age Groups.

Additionally, Smoking and excessive alcohol result in more cases of hypertension.

On the other hand, physical activity significantly reduces the BP to healthy ranges across all groups investigated.

**Click**