**Einleitung**

Der Lungenkrebs ist eine menschenabhängige und auch eine sehr häufig auftretende Krebsart. Daher haben wir den Lungenkrebs auf unserem Projekt bezogen.

Dieser ist ein bösartiger Tumor, der von Bronchialschleimhautzellen oder Lungengewebezellen ausgeht. Ärzte nennen diesen Krebs auch Lungenkarzinom oder Bronchialkarzinom

Das Bronchialkarzinom tritt hauptsächlich im Alter zwischen 50 und 70 Jahren auf. Das durchschnittliche Erkrankungsalter bei Diagnosestellung beträgt nur 69 Jahre.

Im Frühstadium von Lungenkrebs gibt es fast keine typischen Anzeichen und Symptome. Deshalb bleibt es oft lange unentdeckt. Bei Lungenröntgenuntersuchungen finden Ärzte häufig nur zufällig den Lungenkrebs, da der Patient an chronischem Husten, Atemnot oder Kurzatmigkeit leidet.



Abbildung 2

Abbildung 1

Für den Lungenkrebs ist mit weitem Abstand das ***Rauchen*** die Ursache. Ungefähr 90 Prozent aller Bronchialkarzinom-Patienten sind oder waren Raucher. Dabei hängt das Risiko vor allem davon ab, wie viel der Betroffene geraucht hat. Fachleute drücken dies anhand der Packungsmenge aus: Je mehr Zigaretten geraucht wurden, desto höher ist das Krebsrisiko.

Doch nicht nur selber zu rauchen macht krank, auch Passivrauchen erhöht das Risiko, an Lungenkrebs zu erkranken – und zwar um den Faktor 1,3 bis 2. Wer also das Rauchen aufgibt, kann das Risiko für Lungenkrebs deutlich senken. Vor allem für sich selbst, aber auch für seine Mitmenschen.

Wie viele Raucher sterben jährlich am Rauchen?

Rauchen ist das größte vermeidbare Gesundheitsrisiko weltweit. In Deutschland sterben jedes Jahr 110.000 bis 140.000 Menschen an den Folgen des Rauchens. Es gibt mehr als 6 Millionen Menschen auf der Welt. Infolgedessen verursacht Tabakkonsum 15 % der Todesfälle bei Männern und 7 % der Todesfälle bei Frauen.

Wie viel Prozent der Raucher bekommen Lungenkrebs?

Erkranken Raucher häufiger an Lungenkrebs als Nichtraucher? Männer, die rauchen, erkranken 30-mal häufiger an Lungenkrebs als Nichtraucher. Frauen, die rauchen, setzen sich einem 9-fachen Risiko aus. Mit anderen Worten, 91% der Lungenkrebsfälle bei Männern und 65% der Lungenkrebsfälle bei Frauen werden durch Rauchen verursacht. 2013 starben etwa 35.000 Menschen an Lungenkrebs, 91 % waren männliche Raucher, jedes Jahr starben etwa 31.000 Menschen an Lungenkrebs, der mit dem Rauchen in Zusammenhang steht.





Abbildung 3 NICER – Neuerkrankungen; BFS - Sterbefälle

Unsere Aufgabe ist es nun mit einem Datensatz, die oben genannten Fakten bzgl. des Lungenkarzinoms bei ShinyApps zu visualisieren und zu testen.

**Planung des bearbeiteten Datensatzes und der App**

Als wir im Meeting den Datensatz näher betrachten hatten, fanden wir die Variable „Rauchen“ sehr interessant. Denn das Rauchen ist, nehmen wir an, sehr bekannt als die Ursache des Lungenkarzinoms. Das Rauchen kann hier sehr gut in Tabellen und Graphen verwendet werden. Desweiteren, fiel uns auf, dass „Geschlecht“ und „Rauchen“ indentische Werte vorweisten.

Um Ergebnisfehller bzw. Missverständnisse zu vermeiden und eindeutige Ergebnisse visualisieren zu können, haben wir Das Geschlecht von „0“ bzw. „1“ zu „m“ bzw „w“ verändert. Die Liste wurde neu sortiert. Die obligatirische „ID“ haben wir selbstverständlich an erster Stelle, denn so sieht die Liste professionell und auch für den Leser strukturiert aus. „Height und „Weight kommen nacheinander, das diese eine Gruppe aus den Körpermaßen bilden

Neu Hinzugefügte Variabeln

* *Tumour\_size-dichotom* 
  + Der Tumour wurde hier mit T1 und T2 in zwei stufen eingeteilt. Ab einer Größe von 2cm hat der Betroffene die T2
* *Altersgruppen*
  + Durch Altergruppen lassen sich zusätzlich bessere Schlussfolgerungen ziehen, wie z.B bei den Graphen ( Säulendiagramm
* *BMI*
  + Ein erhöhter Cholesterinespiegel kann durch einer fettreichen Ernährung hervorgerufen werden, somit ist der BMI eine wichtige und eine aussagekräftige Variable.
* *Chol\_dichotom*
  + Der Cholesterinespiegel wurde mit niedrig, normal und hoch in drei Stufen eingeteilt. Ab einem Cholwert von 28 ist dieser „Normal“ und ab einem Cholwert von 29 ist dieser “Hoch“.
* *BMI\_dichotom*
  + Der BMI wurde mit „Untegewichtig“, „Normal“ und „Übergewichtig“ in drei Stufen eingeteilt. Unter einer BMI ist die betroffene Person untergewichtig, zwischen 18.5 und 25 normal und ab 25 übergewichtig

Planung von Grafiken und Tabellen

* Grafiken
  + *Scatterplot*

Streudiagramme oder Punktwolken können verwendet werden, um mögliche Beziehungen zwischen zwei Variabeln grafisch darzustellen. Dies ist oft hilfreich, wenn Sie diese Beziehung als Funktion beschreiben möchten. Geplant sind interaktive Scatterplots, in denen man die Variabeln beliebig ändern kann.

* + *Boxplot*

Boxplots, auch Box-Whisker-Plots oder deutsche Kastengrafik genannt, sind Diagramme, das die wichtigsten und zuverlässigsten Positionen und Streuungsmessungen übersichtlich darstellen kann. Dabei werden Minimum, unteres Quartil, Median, oberes Quartil und Maximum dargestellt. Geplant sind hier ebenfalls eine verstellbare bzw interaktive Darstellung mit Boxplots

* + *Säulendiagramm*

Mit Säulendiagrammen werden Datenreihen mittels Säulen grafisch darstellen. Dies ist sowohl mit absoluten als auch mit relativen Häufigkeiten möglich. Verstellbare bzw Interaktive Darstellung ist geplant.

* Tabellen
  + *Kreuztabellen*

In einer Kreuztabelle oder Kontingenztabelle können jeweilige absoluten Häufigkeiten verwendet werden, um statistische Merkmale oder Variablen anzuzeigen (normalerweise nominale oder geordnete Skalenmerkmale, da es normalerweise zu viele Werte gibt, um Merkmale zu messen).Dies ermöglicht die Anzeige und Auswertung von gemeinsamen Verteilungen oder gemeinsamen Häufigkeiten, Randverteilungen oder Randhäufigkeiten und bedingten Verteilungen.

Anschließend ist geplant, mittels Kreuztaballen, Abhängigkeitstest oder Homogenitätstest durchzuführen.

**Übersicht über die Eigenebeteiligung**

**Devrim:**

Durch das Projekt konnte ich einen umfassenden Eindruck gewinnen. Die in den Vorlesungen vermittelten Lerninhalte konnten bei der Ausarbeitung dieses Projektes angemessen angewendet werden und bestätigen den praxisnahen Bezug.

Zum Start der Projektarbeit haben wir die Aufgaben aufgeteilt. Anschließend haben wir nach weiteren Terminen unsere Codes gegenseitig präsentiert. Bei auftretenden Problemen wurden diese analysiert und ausdiskutiert. Die Projektgruppe hat sich mit den Mitgliedern abgestimmt, und die Probleme konnten gelöst werden.

Die Projektgruppe hat die gesamte Projektarbeit in drei großen Themen gegliedert. Ich war zuständig für die Grafiken, wobei bei Problemen, die auftauchten, andere Mitglieder mir geholfen haben. Das Gleiche gilt auch andersrum. Ich habe noch dazu beigetragen das Projekt etwas interaktiver zu gestalten (zb. html usw.), und die App mit Informationen (About) ergänzt.

Während des Projektes konnten verschiedene Eigenschaften, wie zum Beispiel, die Organisation von Besprechungen und Umgang mit plötzlich auftretenden Problemen erlernt werden.

<https://www.massmatics.de/merkzettel/#!963:Scatterplot>

<https://studyflix.de/statistik/boxplot-1044>

https://studyflix.de/statistik/histogramm-1711