**Praktikum – Infos über den Datensatz**

AUFGABENSTELLUNG :

Ziel: Analyse der Trends im Substanzkonsum, Bestimmung des Bedarfs an Präventions- und Behandlungsprogrammen

* Projekt soll sich mit der Untersuchung von Zusammenhängen zwischen Drogenkonsum, demografischen Faktoren und mentaler Gesundheit beschäftigen

Fragenstellung:

* Wie hängt der Konsum im Jahr 2019 mit demographischen Merkmalen zusammen?
* Wie haben sich die Konsummuster in dem Zeitraum von 2015 bis 2019 entwickelt?
* Wie steht der Konsum im Jahr 2019 mit dem psychischen Gesundheitszustand der Befragten in Verbindung?

(multiyear trend analysis)

*Was ist der NSDUH Datensatz?*

* Jährliche Befragung amerikanischer Bürger ab dem Alter von 12
* Hauptquelle über statistische Daten über den Gebrauch von Tabak, Alkohol, pharmazeutische verschriebene Drogen (Schmerzmittel, , Aufputschmittel und Beruhigungsmittel (Tranquilizer und Sedatives) ) und Drogen wie z.B. Marihuana, Kokain
* Der Datensatz enthält zudem Infos zu SUDs, das Treatment von Abhängigkeit und zu mentalen Gesundheitsproblem und deren Behandlung
* Datenschutz: /
* Zum Format: gleiche Variablen sind zusammen im Datenset angeordnet (auf die „Originale“ Variable folgt dann die „Imputed“ oder „recoded“ Version) -> erkennbar an Klammern! 🡪 enthalten of missing data
  + Bei diesen Variablen wird empfohlen Schätzwerte zu verwendet statt roh- oder editierten Daten
  + Sortierung allgemein: erst allgemeine Daten (Fallspezifizierung, geographische Daten, Werte zu Masse, ...) dann folgen die demografischen Werte (Bildung, Arbeit, Krankenversicherung, Gehalt …) , dann editierte Daten über spezifische Drogen, dann Daten über physische und mentale Gesundheit, abschließend Daten zum Haushalt „roster“ (?) und Infos über die Interviews an sich

*Wie läuft die Umfrage ab?*

* Die 8 Staaten mit der größten Population (insgesamt 48% der gesamten GG) gelten als großen „Sample states“ mit einer Stichprobengröße von bis zu 3600, in den anderen Staaten wurden bis zu jeweils 900 UE befragt
* 2015: Verteilung der Altersgruppen: 25% Jugendliche, 25 junge Erwachsene (18-25), 15 % für 26-34 Jährige, 20% für 35-49, 15% für Erwachsene älter als 50
  + Fragen über Schnupftabak wurden ersetzt mit Fragen über tabakfreie Zigaretten usw
* Some questions are interviewer administered, some self administered
* Befragten wurden nur einmal befragt 🡪 cross-sectional Survey

Strategie und Herangehensweise:

* In jedem Staat werden SSR geformt (state sampling regions), die jeweils geografisch gleich groß sind 🡪 gleiche Probengröße sollte damit erreicht werden
* Insgesamt gab es 750 solcher SSRs
* Die erste Auswahl der Befragten erfolgt durch einen Census (diese Phase beginnt mit einer einer Konstruktions von einer sample größe so dass jeder census ein „Record“ enthält

*Stärken und Schwächen der Umfrage*

Gut:

* Fragen sind bei ACASI registriert 🡪 garantiert hohen Datenschutz und Anonymität 🡪 mehr ehrliche Antworten
* Proben sehr groß und stark verteilt 🡪 man kann auch Aussagen für mehr Bezirke treffen auch auf substate areas

Schlecht

* Daten beruhen auf Eigenberichte der Befragten -> Erinnerung? Ehrlichkeit?
* Underreporting und overreporting
* Fragen zu Medikamenten: Unklar auf wlechen Stoff genau man reagiert/ süchtig sein könnte
* 3% der Bevölkerung ist aus der GG ausgeschlossen 🡪 könnte dazu führen, dass die Schätzer für z.B. allgemein verbreitetet und benutzte Drogen (Heroin) nicht akkurat sind

Diese Substanzen Betrachten wir genauer

Tabak

Alkohol

Kokain

Heroin