**Was wurde gemacht und wie ist der aktuelle Zustand?**

* Beschaffung von Datensätzen vom statistischen Bundesamt (Tabellen zur Unfallstatistik in 2019 und 2020, Bevölkerungsentwicklung der letzten 10 Jahre, Einwohnerzahlen in den einzelnen Landkreisen, ÖPV Daten aus 2019 und 2020) und die Anpassung dieser Daten zum gewünschten Format
* Aufbereitung der RKI-Daten (unnötige Spalten entfernen; Datentypen anpassen, Landkreis\_ID in numeric)
* Vereinheitlichung der Daten (zB Landkreisnamen in Bevölkerungs- und RKI Tabelle)
* Visualisierung der einzelnen Landkreise (Vergleich), Altersgruppen, Korrelation der Bevölkerungsdichte mit der 7-Tage-Inzidenz.
* Wir stellen die Hypothese auf, das Corona anfangs andere Bevölkerungsschichten erreicht hat als später. Zu Beginn waren hauptsächlich Skifahrer und Reisende betroffen, Menschen mit potenziell mehr Einkommen (so die Hypothese). Wir vermuten auch, dass nun ärmere Schichten betroffen sind, da diese gegebenenfalls nicht so viel Vitamin B bei der Impfterminbesorgung hatten.
* Zur Untersuchung dieser Hypothese berechnen wir das Arbeitsentgelt pro Kopf nach Landkreis. Anschließend wird die Korrelation zwischen Inzidenz und Einkommen abhängig von der Kalenderwoche visualisiert.
* Anteil der Fälle mit bekanntem Erkrankungsdatum im Verlauf der Pandemie und Verteilung der Differenz zwischen Erkrankungs- und Meldedatum.
* Vergleich der ÖPV Personenkilometer im Jahr 2020 mit denen im Jahr 2019. Zusätzlich wird die Unfallstatistik zum Vergleich herangezogen. Hierbei fällt auf, dass sowohl die Anzahl der gefahrenen Personenkilometer im ÖPV als auch die der Unfälle im Jahr 2020 im Vergleich zum vorherigen Jahr stark zurück gegangen sind.

Der Zustand des Projektes ist momentan noch mehr eine Ansammlung von funktionierenden R-Skripten anstatt eines vollkommen funktionstüchtigen R-Paketes. Wir werden uns später noch damit beschäftigen müssen, die Nomenklatur und Funktionsweise der einzelnen Skripte zu vereinheitlichen, damit daraus ein R-Paket entsteht. Damit ist der Plan für die nächsten Wochen,

den Inhalt abzuschließen und zu deuten und dann das Paket zu formen.

**Was ist noch geplant?**

* Die Skripte zu einem R Package zusammenfassen
* Daten zu täglichen Impfungen in Deutschland besorgen, um den Effekt der Impfkampagne quantifizieren, wobei man sich noch überlegen muss, wie man die Impfeffekte und andere Einflüsse auf die Inzidenz (zB saisonale) separieren kann, um ein klare Aussage zu treffen
* RKI-Tabelle auf Konformität mit Tidy R Standards überprüfen und evtl. anpassen
* Vergleich Todessszahlen in 2020 insgesamt mit Jahren davor nach Alter