

Logboek

Dinsdag 5 juni:

- Github laten werken
- Data ingeladen en bekeken
- Begin gemaakt data cleaning

Woensdag 6 juni:

- Drive bestanden opgezet met planning, eerste ideeën mogelijke grafieken
- Data cleaning afgemaakt
 - NA's toegevoegd op witregels
 - Missende adressen aangevuld met behulp van Geo-coder
- Eerste resultaten gevonden

Donderdag 7 juni:

- Meeting met Matthijs

Vrijdag 8 juni & Maandag 11 juni:

- Thuis aan eigen deel gewerkt
 - Tessel: scatterplot over kaart van Amerika
 - Jeroen: lat/long gebruiken om adressen te vinden
 - Raymon: grote verschillen weergeven per staat/jaar
 - Chang: relatie tussen killer en victim bekijken

Dinsdag 12 juni:

- Meeting met Matthijs
- Eerste deel grafieken plotten

Dinsdag 19 juni:

- Iedereen: Meeting met Matthijs
- Tessel: Geomap afgemaakt
- Jeroen: Geomap afgemaakt
- Raymon: 1 grafiek gemaakt en 2 afgemaakt
- Chang: begin aan website en een grafiek gemaakt.

Processbook

Dinsdag 19 juni:

Bij het schrijven van het stuk code voor de plot 'murders on map' zijn wij tegen een aantal problemen aangelopen, die we op verschillende manieren hebben opgelost. Om te beginnen hadden wij wat problemen met het genereren van een passend kleurenpalet om de gradatie van de gevaarlijkheden van de staten aan te geven. Wij wilde in eerste instantie alle 46 gebruikte staten een apart kleur geven, en deze dan laten oplopen van minst gevaarlijk naar meest gevaarlijk. Maar het bleek lastig te zijn om zelf een specifiek aantal oplopende kleuren te genereren. Wij hebben er daarom voor gekozen om het standaard palette Viridis256 te

gebruiken. Als gevolg hebben niet alle staten een verschillende kleur, maar er is wel duidelijk te zien hoe de gevaarlijkheid van de staten oploopt.

Daarnaast hebben wij de kleuren van Viridis256 omgedraaid, om een duidelijkere plot te creëren. Als gevolg hiervan moest ook de legenda worden omgedraaid. Hier hebben wij flink voor gezocht op het internet, maar geen goede oplossing voor gevonden. Daarom hebben wij gekozen om de getallen bij de colorbar weg te laten, en zelf een tekstvak in te voegen met hierin de getallen die bij de verschillende kleuren horen.

Woensdag 20 juni:

Vandaag was er een probleem met het gebruiken van de github. Jeroen wilde een groot en belangrijk stuk code voor het maken van de scatterplot op de kaart van Amerika pushen, maar kreeg toen een aantal errors. Uiteindelijk kreeg hij de oude versie terug, en waren alle vorderingen van dinsdag verdwenen. Wij zijn dus de rest van de dag bezig geweest om de code opnieuw te schrijven en de schade te beperken.

Donderdag 21 juni:

Chang heeft de opzet voor de website verder voltooid. Tessel is begonnen met het schrijven van het verslag. Raymon is verder gegaan met het plotten van grafieken. Jeroen heeft de scatterplot over de kaart van Amerika verder verbeterd.

Maandag 25 juni:

Tessel heeft geprobeerd om een hoovertool te maken voor de kaart van Amerika. Na veel onderzoek bleek het echter erg ingewikkeld om zelf kolommen uit de dataset in te geven als tooltips. Na een aantal uur hieraan te hebben gewerkt is besloten om de focus te verleggen en verder te werken aan het verslag. Chang had wat problemen met het centreren van de plots op de website, maar door de hulp van Mattijs is dit opgelost. Jeroen is aan de slag gegaan met het zoeken van andere datasets om een correlatie te vinden met onze eigen dataset. Er is gekozen om te focussen op het gemiddelde inkomen van de inwoners van een staat, en dit te vergelijken met het aantal incidenten. Wij verwachten dat een grote inkomensongelijkheid gecorreleerd is aan een hoger aantal incidenten.

Dinsdag 26 Juni

Chang heeft lineaire regressie toegepast op de dataset. Bij de uitkomst heeft hij één resultaat weggelaten, omdat deze in conflict was met de rest van de resultaten. Jeroen en Raymon zijn weer aan de slag gegaan met de hoovertool van de Geomap. Na een dag hieraan te werken is dit uiteindelijk toch gelukt. Ondertussen is Chang bezig geweest met het ontwerp van de website en deze zo mooi en strak mogelijk te krijgen.

Tessel is verder gegaan met het schrijven van het verslag. Er is gekozen om hier een iemand op te laten focussen, omdat wij uit ervaring hebben gemerkt dat het inefficiënt is om met meerdere mensen aan een verslag te werken. Tijdens het schrijven van het verslag is natuurlijk wel regelmatig overleg geweest met de rest van de groepsleden.

Woensdag 27 juni:

Jeroen en Raymon hebben een extra kaart gemaakt waarbij de poverty rate van de staten vergeleken wordt met het aantal incidenten. Er werd namelijk een correlatie tussen een hoge poverty rate en veel incidenten verwacht. Verder zijn we de hele dag bezig geweest met het verbeteren van de opmaak en interactiviteit van de overige grafieken.

Er is een kaart gemaakt waarin de relatieve gevaarlijkheid van de staten wordt weergegeven. Hierbij is het aantal incidenten gedeeld door de populatie, als resultaat is er dus een kaart gemaakt waarbij het aantal gewonden en doden per hoofd van de bevolking is weergegeven. Hierbij is de District of Columbia buiten beschouwing gelaten voor het toedelen van scores aan de staten. Deze staat is relatief zo gevaarlijk, dat de rest van de staten anders geel kleurden, en er geen verschil te zien was tussen de andere staten. Vandaar dat ervoor is gekozen om deze staat buiten beschouwing te laten.

Donderdag 28 juni:

Het eerste deel van de dag hebben wij allemaal ons deel van het project afgerond. Raymon en Jeroen hebben de laatste grafieken afgemaakt, Chang heeft de laatste hand gelegd aan de website en Tessel heeft het verslag zo goed als af gemaakt. Het tweede deel van de dag hebben wij gebruikt om samen het verslag door te nemen, en de presentatie van vrijdag voor te bereiden.