



PYTHON PROGRAMMING

Python Fundamentals

Naam: Stefan Schurgers
Studentennummer: 900102969

Inhoudsopgave

Plan van aanpak.....	2
Verantwoording.....	3
Beschrijving en ontwerp van de functie analyse:.....	3
Ontwerp 2 zoekfuncties:.....	3
Toelichting gemaakt keuzes.....	3
Testverslag.....	4
Projectevaluatie.....	4

Plan van aanpak

Opzet van menu keuze:

1. De bestanden automatisch laten inlezen, dit voorkomt onnodige if else statements in menu keuzes waarbij data moet worden getoond of worden aangepast.
2. Maak een Class voor de volgende objecten: Bedrijven, Rapporten, Inspecteurs en gassen. Onder deze Class objecten worden verschillende functies aangemaakt.

	Class Object	Functies binnen Class object
Bedrijven		lees_bedrijven() toon_bedrijven() zoek_locatie() zoek_naam() wijzigen_bedrijf() toevoegen_bedrijf() bepalen_boetes() analyse()
Rapporten		lees_bedrijven() toon_rapporten_inspecteurs() toon_rapporten_bedrijf() wijzigen_rapport() toevoegen_rapport()
Inspecteurs		lees_bedrijven() toon_inspecteur()
Gassen		lees_gassen() toon_gassen()
Gassen_plot		plot_gas()

3. Deze verschillende classes worden opgeslagen in verschillende python bestanden. Deze worden dat met de import functie ingelezen en toegepast in het menu bestand.

Het aantal uren dat ik wil besteden aan de opdracht zijn 8 uur per week. Helaas is dit doel niet behaald. Ik heb vooral in de laatste 2 weken aan de opdracht gewerkt. Maar dit was wel gemiddeld 6 uur per dag.

Verantwoording

Beschrijving en ontwerp van de functie analyse:

Op basis van opdracht 7 van de eindopdracht wordt een dataframe gegenereerd met de bekende bedrijven en de bijbehorende omgevingen. Dit gebeurt door de coördinaten van het bedrijven bestand te gebruiken in combinatie met een for loop en een range. Zo ontstaat er een lijst met coördinaten die binnen het 5 maal 5 kwadrant vallen.

Dit bestand bevat dus ook de coördinaten van de bekende bedrijven. Vervolgens wordt er een anti join gemaakt op het gassen bestand. Zo blijft er van de gassen dataset alleen data over van niet bekende bedrijven. Vervolgens wordt de dataset gesorteerd van groot naar klein op basis van de totale uitstoot berekend in opgave 5 "Inlezen metingen bestand". Zo ontstaat een overzicht met de uitstoot en de daarbij behorende coördinaten.

Ontwerp 2 zoekfuncties:

De gebruiker wordt gevraagd om de x en y coördinaten in te vullen van een desbetreffend bedrijf. Deze ingegeven waarden van de x en y coördinaten worden van de daadwerkelijke x en y coördinaten afgetrokken. Het verschil van de daadwerkelijke x en y coördinaten worden bij elkaar opgeteld. Vervolgens wordt er gesorteerd van klein naar groot. Dan wordt er een subset van de bedrijven dataset genomen waarin alleen de eerste regel overblijft. Dus de waarde met het kleinste verschil in afstand.

Voor de zoekfunctie met deels bedrijfsnaam wordt de gebruiker gevraagd om een naam in te voeren. Vervolgens wordt `str.contains()` in combinatie met de input van de gebruiker om de bedrijfsnaam te filteren. Hier wordt ook weer een subset genomen van de eerste regel zodat het eerst gevonden bedrijf wordt weergegeven.

Toelichting gemaakt keuzes

Voor deze opdracht heb ik voornamelijk gebruik gemaakt van de pandas bibliotheek. Gezien de applicatie gebaseerd is op data en pandas uitermate geschikt is om data in te lezen, data te manipuleren en te analyseren heb ik gebruik gemaakt van deze bibliotheek. Echter is deze niet in de les behandeld dus heb ik wel extra onderzoek werk moeten verrichten. Daarnaast heb ik voor de visualisatie gebruik gemaakt van matplotlib en seaborn. Seaborn is geschikter om pandas dataframes te visualiseren. Daarom was ik wel in zekere mate genooddaakt om ook seaborn te gebruiken in combinatie met matplotlib.

Daarnaast heb ik vooral van classes gebruik gemaakt in verschillende python bestanden. Zo kon ik kleinschalig iets programmeren en testen. Als er iets fout gaat is het makkelijker te lokaliseren waar het probleem precies ligt. Buiten dat het een vereiste was om met modules te werken is het ook een handige methode om meerdere functies te verwerken. Ook dit geeft overzicht en structuur.

Testverslag

De afzonderlijke methodes die zijn opgesteld binnen de gedefinieerde classes zijn afzonderlijk gedraaid. Daarnaast is er ook getest op foutieve ingegeven data bij input functies die zijn gebruikt binnen de verschillende methoden om te testen of de programmatuur niet stopt of een errors weergeeft. Indien er foute gegevens worden ingevoerd worden deze opgevangen door een while loop te gebruiken.

Gedurende het testen is me opgevallen dat wanneer de from import functie wordt gebruikt om de modules te laden er in de prompt gevraagd wordt naar de inhoud van input functies die hierin zijn gedefinieerd. Ik vind het erg opvallend want als ik de Class draai binnen het desbetreffende python bestand wordt dit niet gevraagd. In alle eerlijkheid heb ik dit probleem niet kunnen oplossen tijdens de testen. Dus een gebruiker die de modules importeert wordt gelijk om input gevraagd wordt dit niet geleverd crasht het systeem. Binnen de lesstof heb ik geen oplossing gezien voor dit probleem. Hierdoor kon ik alleen de bestanden apart testen zonder deze te testen in zij geheel in het keuze bestand.

Projectevaluatie

Over de opdracht heb ik wel gemengde gevoelens. Ik heb wel geleerd om met classes/modules te werken alsmede deze over meerdere python bestanden te verspreiden en toe te importeren en toe te passen. Echter is de moeilijkheidsgraad wel erg hoog bij een aantal opdrachten. Met name opdracht 7 vond ik erg lastig. Nu heb ik wel met pandas gewerkt en ik weet dat dit geen onderdeel is van de leerlijn. Maar soms zag ik echt geen andere uitweg dan deze bibliotheek te gebruiken. Zeker gezien er informatie van een dataset naar de andere dataset moest worden omgezet. Zonder gebruik te maken van pd.merge en puur alleen uitgaan van de lesstof zou mij dit gewoonweg niet zijn gelukt.

Daarnaast vond ik de opdracht niet een realistische applicatie die in het bedrijfsleven zou worden ontwikkeld. De opdracht is wel geschikt om te oefenen met python maar wellicht is het een idee om een data applicatie met dash te bouwen. Dit is een visualisatie bibliotheek waarmee dashboards gegenereerd kunnen worden, iets wat in de praktijk vaker gedaan word en ook python mee geleerd kan worden. Ik denk dat dit de motivatie wellicht kan bevorderen omdat er een applicatie gemaakt wordt die ook in het werkveld gebruikt kan worden.

Echter vind ik het erg frustrerend dat ik een applicatie inlever waarvan ik de afzonderlijke modules niet heb kunnen testen in het keuze bestand. Voor mijn gevoel ben ik erg dichtbij een eind oplossing maar heb ik helaas niet de tijd gehad om een oplossing te vinden voor het probleem. Wellicht had ik alle input vragen bij het keuze menu bestand moeten zetten. Maar zoals gezegd had ik de tijd niet meer dit aan te passen.

Ik moet wel toegeven dat er veel wordt gevraagd bij de opdracht en aan de ene kant is dat wel goed dat je getest wordt om je python kennis toe te passen. En soms is dat wel frustrerend maar in die zin

heb ik wel veel geleerd over python en panda's, al was dat laatste niet de bedoeling van deze leerlijn vind ik het wel fijn om mee te nemen.

Samenvattend vond ik de opdracht wel uitdagend op python niveau op een aantal momenten frustrerend en niet perse praktisch voor het werkveld.