

# Logboek

Datum	Persoon	Wat is er gedaan?
05-06-2018	Louis	Vorbereiden data processing
05-06-2018	Karst	programma maken voor data uit de site halen
05-06-2018	Franklin	Data uit json in dataframe, maken van DataNotebook.
05-06-2018	Remco	Begin gemaakt met ophalen JSON van site.
06-06-2018	Louis	Gekeken naar JupyterHub. Meegeholpen met missing data.
06-06-2018	Karst	grafiekengenerator gemaakt voor exploratory data analysis
06-06-2018	Franklin	Data verder geanalyseerd en cleanen, begrijpen van de data. Mysterie van de ' <i>missing data</i> ' opgelost.
06-06-2018	Remco	Descriptions, attribute types, missing values bekeken. Begrijpen van de data.
07-06-2018	Louis	Data cleanen
07-06-2018	Karst	grafiekengenerator gebruikt nu de goede data en slaat de grafieken op.
07-06-2018	Franklin	Data verwerkt, meegewerkt aan het cleanen af te ronden van de data.
07-06-2018	Remco	Data analyse en verder cleanen.
08-06-2018	Louis	
08-06-2018	Karst	
08-06-2018	Franklin	EDA bestuderen, pdf doorgenomen en online tutorials over eda.
08-06-2018	Remco	Data cleanen afgerond en nieuwe data frames in .csv bestanden gezet
11-06-2018	Louis	Data-analyse
11-06-2018	Karst	Code om gemiddelde prijs van alle bedrijfstakken te vinden gemaakt.
11-06-2018	Franklin	Nieuwe databronnen, research voor eindproduct. Voorstel opgesteld en gepresenteerd. Bij Karst meegekeken hoe de grafieken worden gemaakt.

11-06-2018	Remco	Begin data-analyse
12-06-2018	Louis	histogram startdatums gemaakt
12-06-2018	Karst	
12-06-2018	Franklin	
12-06-2018	Remco	Descriptions in .txt gezet, Afzet.csv gemaakt en Bedrijfstakken.csv gemaakt
13-06-2018	Louis	Data cleanen
13-06-2018	Karst	
13-06-2018	Franklin	Onderzoek naar manieren om nieuwe features te maken. Hypotheses opstellen. Nieuwe data bronnen gevonden.
13-06-2018	Remco	Data-analyse en verder cleanen
14-06-2018	Louis	
14-06-2018	Karst	heatmap van de correlatie tussen alle bedrijven gemaakt.
14-06-2018	Franklin	Nieuwe data kolommen gemaakt, heatmap gemaakt, data weer onder de loep genomen en bruikbaarheid gevonden, bugs fixen.
14-06-2018	Remco	Data-analyse
18-06-2018	Louis	Overkoepelende takken onderzocht en geanalyseerd
18-06-2018	Karst	
18-06-2018	Franklin	Feature engineering
18-06-2018	Remco	Begonnen met rapport maken
19-06-2018	Louis	Dataframe voor features gemaakt
19-06-2018	Karst	
19-06-2018	Franklin	Nieuwe features gemaakt waaronder alleen binnenlands, percentages van wegingscoëfficiënt tov branche.
19-06-2018	Remco	Vragen beantwoorden, relaties zoeken en bronnen zoeken
20-06-2018	Louis	
20-06-2018	Karst	

20-06-2018	Franklin	Verder pogen nieuwe dimensies aan dataset toe te voegen vanaf periode 2005. Scatter plots maken. Clusteren gaat niet bij deze dataset. Kijken hoe we ML kunnen toepassen op deze dataset.
20-06-2018	Remco	Analyse
21-06-2018	Louis	Dataframe van overkoepelende takken gemaakt
21-06-2018	Karst	
21-06-2018	Franklin	Taakverdeling opstellen en eind deliverables voorbereiden. Product Demo verdelen en structuur geven.
21-06-2018	Remco	Inleiding, methode begonnen
25-06-2018	Louis	Github pages
25-06-2018	Karst	dip.py has made more graphs and now uses preexisting graphs instead of generating new ones.
25-06-2018	Franklin	Regressie toepassen totaal en overkoepelend binnenlands/buitenlands. Beginnen aan presentatie verhaal.
25-06-2018	Remco	Discussie begonnen
26-06-2018	Louis	Github pages
26-06-2018	Karst	
26-06-2018	Franklin	Product demo af.
26-06-2018	Remco	Rapport afgemaakt
27-06-2018	Louis	Rapport aangevuld
27-06-2018	Karst	
27-06-2018	Franklin	
27-06-2018	Remco	
28-06-2018	Louis	Website aangepast
28-06-2018	Karst	
28-06-2018	Franklin	
28-06-2018	Remco	
29-06-2018	Louis	Presentatie

29-06-2018	Karst	Presentatie
29-06-2018	Franklin	Presentatie
29-06-2018	Remco	Presentatie