**Opzet Python-cursus 25-11-2019**

Welkom

Controle apparatuur

* Welke OS-en;
* Welke Python versies;
* Virtual environments per OS nagaan;
* IDE’s?

Herhaling

* Inspringen
* Taal elementen:
  + Variablelen
    - Naamgeving (geen camel-case)
  + If / Elif
  + While
  + For
  + Open
  + Input
  + With (context manager)
  + Oefeningen
  + Data-typen
    - String
    - Set
    - Tuple
    - List
    - Dictionary
  + Comprehensions
    - List comprehension
    - Dict comprehension
  + Oefeningen
* Virtual environments
* Imports (packages)

Pause?

Hoofd gedeelte

Analyzeren van de 2019 Stack Overflow Developer Survey

Vraag: “Language popularity by Developer type”. Of: welke computertalen worden veel gebruikt wanneer embedded computers (micro controllers) worden gebruikt?

Website:

<https://insights.stackoverflow.com/survey/2019?utm_source=so-owned&utm_medium=blog&utm_campaign=dev-survey-2019&utm_content=launch-blog>

Data:

<https://insights.stackoverflow.com/survey>

Opzet

* Aanmaken virtual environment
  + mkdir cursus-2019-11
  + cd cursus-2019-11
  + python -m venv venv
  + Bekijken venv/ directory
  + source venv/bin/activate
  + which python3
  + which pip
* Start de IDE
  + code .
  + Korte uitleg van Visual Studio Code
    - Denk ook aan python/lint extensions

* Downloaden CSV handmatig of via Python? Toch even leuk om een data download te doen via Requests.
  + Zie source.
* ZIP uitpakken en plaatsen in Data directory
* Data bekijken, bv Readme.txt, data en schema.
* Maak een sof-analyze.py
* Daarna:
  + import CSV
  + Open het bestand:
    - with open('data/survey\_results\_public.csv') as f:

csv\_reader = csv.DictReader(f)

* + Print daarna de eerste regel:
    - for line in csv\_reader:
    - print(line)
    - break
  + Laat zien dat wat je ziet een dictionary is (van tuples)
  + Laat zien dat je een alleen veld kunt printen door een key aan te geven:
    - print(line[“Hobbyist”])
  + Verwijder eventueel het break-statement.
  + Percentage hobby-programmeurs? 3 Manieren:
    - Standaard manier
      * Variabelen voor ‘yes’ en ‘no’ plus increment ervan
      * Ná het with-statement:
        + plaats yes\_counter en no\_counter
        + Increment counters als line[“Hobbyist”] == “Yes” / “No”
        + Maak een total counter en yes\_pct (yes\_counter / total \* 100)
        + Voer een round door
        + print resultaten
        + Laat line.get(“Hobbyist) zien + default waarden.
    - Gebruik van een dictionary (histogram?)
      * Verwijder yes- en no\_counters
      * Maak een counter dict:
        + counter = {“Yes”: 0, “No: 0}
      * Het If/Elif commando kan nu vervangen worden door:
        + counter[“Hobbyist”] += 1
        + Wijzig nu ook alle andere variabelen en de print statements
      * Leg nu het gebruik van defaultdict uit (in REPL).
        + From collections import defaultdict
        + Wijzig de counter in: counter = defaultdict(int)
    - Via Counter-object in Collections
      * From collections import Counter
      * Leg Counter uit, bijvoorbeeld in de REPL
      * Wijzig de counter in: counter = Count()
  + Nu over naar: Wat zijn de meest populaire computertalen?
    - Open de file schema.csv en ga na in welke kolom de computertalen staan: LanguageWorkedWith (AQ)
      * Wijzig de counter in ‘language\_counter’
      * print de eerste regel met ‘language\_counter’
      * Laat het type zien van ‘language\_counter’ (str)
      * Laat eventueel de eerste 10 regels zien mbv enumerate en If i < 11
    - Een regel met languages bevat meerdere computertalen: gescheiden met ‘;’
      * Lees een LanguageWorkedWith regel in en split deze op ‘;’ (languages\_list)
      * Print het resultaat
      * Doe het eventueel voor in de REPL
    - Parse de language\_list met een for-loop
      * for language in language\_list:
    - Tel de talen: language\_counter[language] += 1
    - print de language\_counter
    - De for-loop kan vervangen worden door: language\_counter.update(languages\_list)
    - Doe het voor in de REPL
    - vervang de loop en print de language\_counter
    - Laat deze eventueel over de eerste 10 regels zien.
    - Verwijder nu de regel restrictie en verwerk alle regels.
    - Wijzig het print statement in language\_counter.most\_common(5) om de 5 meest gebruikte computertalen te laten zien.
    - Laat zien dat de output van most\_common een list met tupels is.
    - Maak een total counter en tel het aantal regels.
    - Ga nu met een for-loop alle talen en hoe vaak deze worden gebruikt na:
      * for language, value in language\_counter.most\_common(5):
        + language\_pct = round((value / total) \* 100, 2)
        + print(f"{language.rjust(10)}: {language\_pct}%")
  + Nu naar de vraag: “Talen per Developer Type”
    - Laat zien dat in de data het veld DevType alle developertypes bevat. Gescheiden door ‘;’
    - Verwijder total en language\_counter. Maak een dev\_type\_info als een lege dictionary: dev\_type\_info = {}.

Verwijder verder alles behalve de for-loop over de lines.

* + - Dan, voor iedere line:
      * dev\_types = line[“DevType”].split(“;”)
    - Ga nu bij alle dev\_types langs en maak een lege dict voor iedere dev\_type key:
      * for dev\_type in dev\_types:
        + dev\_type\_info[dev\_type] = {}
        + dit levert: {'Designer': {}, 'Developer, back-end': {}, 'Developer, front-end': {}, 'Developer, full-stack': {}}
        + Laat dit in de REPL zien.
    - Print de verschillende dev\_type keys:
      * for key in dev\_types\_info:
        + print(key)