01

OVERVIEW OF HOMEWORK 1B

HW 1b - Extract & Load

DEPLOY A TEMPLATE

(Every Homework)

01 Open Azure CloudShell

02 Clone Github Repository

03 Create Azure Template

04 Deploy Azure Template

slidesmania.com

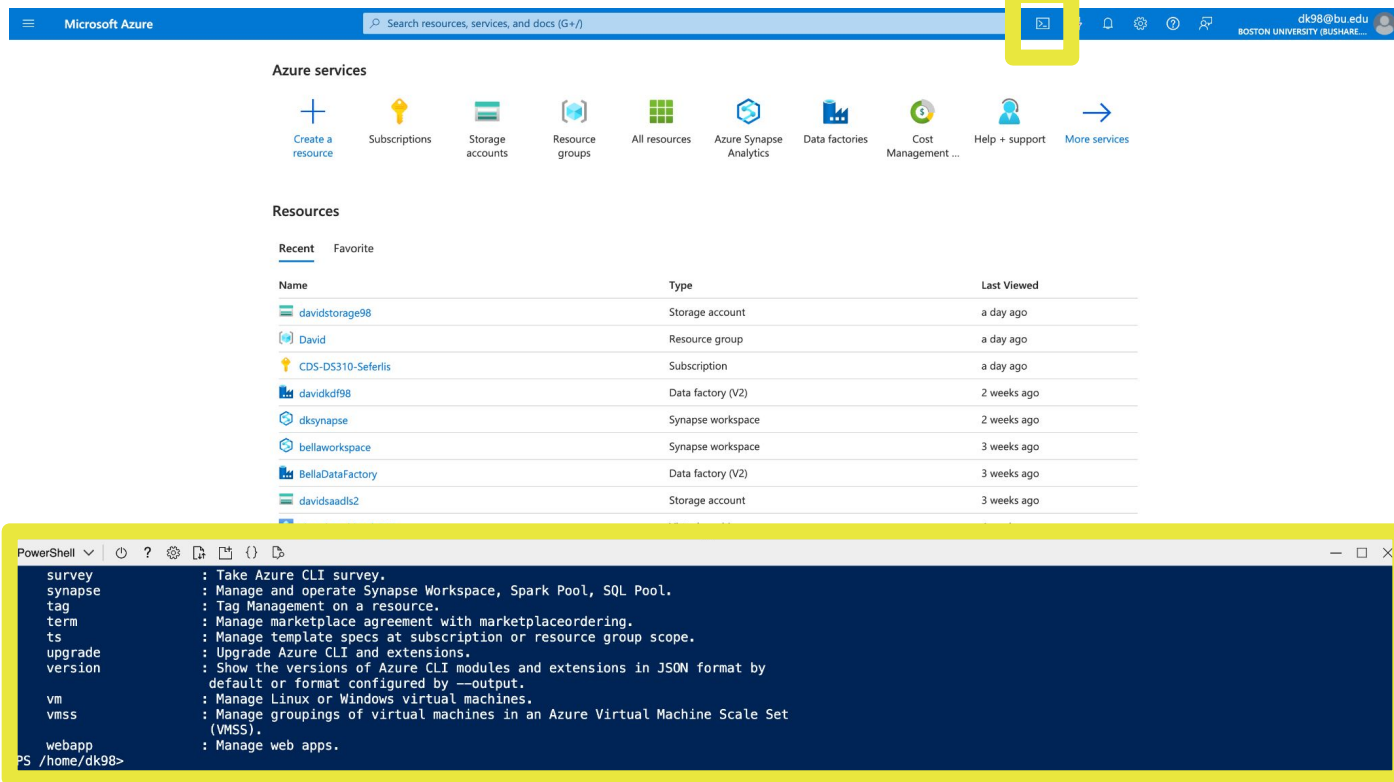# WHY?

$100 Free Credit is not a lot, and your **custom resource settings** may drain it even faster!

1. OPEN AZ CLOUD SHELL

AZURE CLOUD SHELL

- **Magic Shell**

    - Can **ask the sea** to do things like bring you a wave, calm down, or even change the tide.

- **Cloud Shell**

    - Can **ask the cloud** to start services, manage them, or even troubleshoot issues, just by typing in commands.
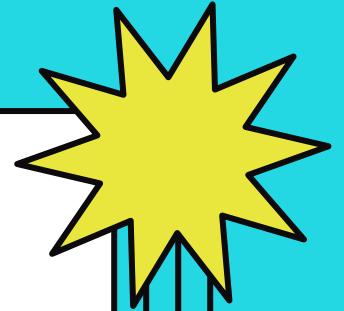
LINUX COMMAND

- **cd (Change Directory)**
  - change the current working directory

- **ls (List)**
  - list the contents of a directory

Downloads

Name
> 📁 untitled folder
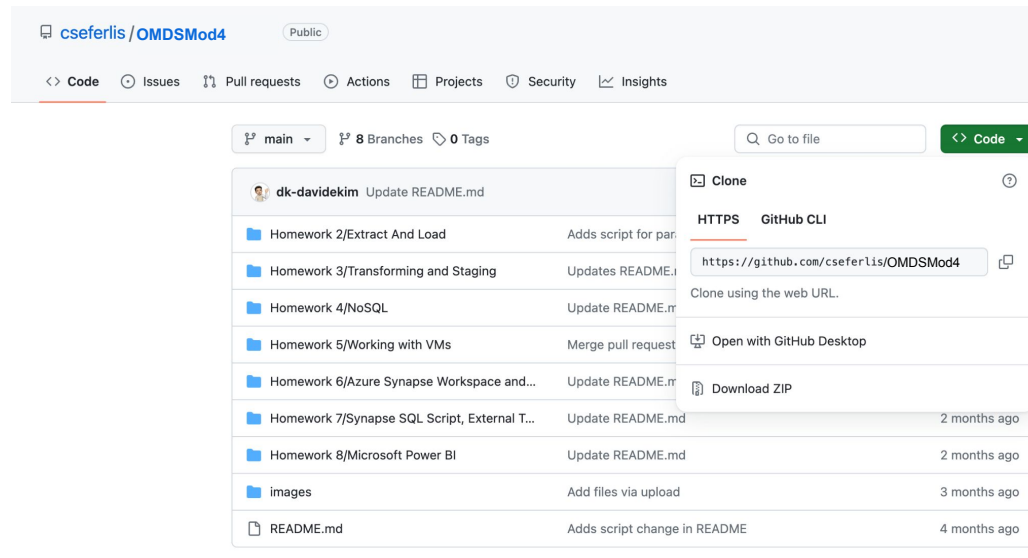
```
~/Downloads (0.028s)
ls
untitled folder
```

```
~/Downloads (0.022s)
cd 'untitled folder'

~/Downloads/untitled folder
```

untitled folder

Name
📄 helloDS310.txt

```
~/Downloads/untitled folder
ls
helloDS310.txt
```
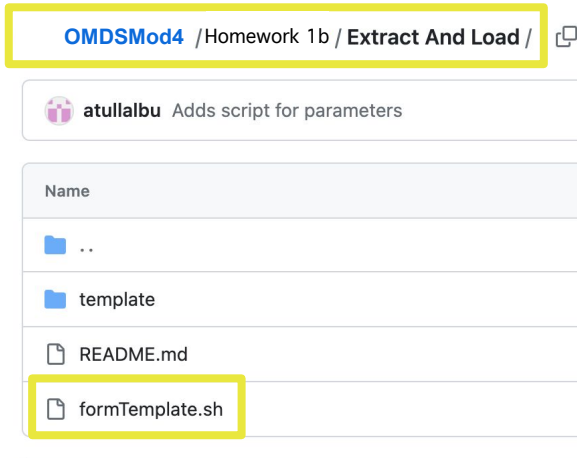
2. CLONE GITHUB REPO

cseferlis / **OMDSMod4**   Public

<> Code   Issues   Pull requests   Actions   Projects   Security   Insights

main   8 Branches   0 Tags   Go to file   <> Code

dk-davidekim Update README.md   Clone   ?

Homework 2/Extract And Load   Adds script for par...   HTTPS   GitHub CLI

Homework 3/Transforming and Staging   Updates README.r

Homework 4/NoSQL   Update README.m   https://github.com/cseferlis/OMDSMod4

Homework 5/Working with VMs   Merge pull request   Clone using the web URL.

Homework 6/Azure Synapse Workspace and... Update README.m   Open with GitHub Desktop

Homework 7/Synapse SQL Script, External T... Update README.md   2 months ago Download ZIP

Homework 8/Microsoft Power BI   Update README.md   2 months ago

images   Add files via upload   3 months ago

README.md   Adds script change in README   4 months ago

https://github.com/cseferlis/OMDSMod4.git

```
PS /home/dk98> git clone https://github.com/cseferlis/ OMDSMod4 .git
```

3. NAVIGATE TO FORMTEMPLATE.SH

OMDSMod4 / Homework 1b / **Extract And Load** /

atullalbu Adds script for parameters

Name

..

template

README.md

formTemplate.sh

```
PS /home/seferlis/OMDSMod4/Homework 1b/Extract And Load>
```

```
bash ./formTemplate.sh
```

Template and parameters created successfully: ./template/template.json and ./template/parameters.json

4. DEPLOY AZURE TEMPLATE

`Template and parameters created successfully: ./template/template.json and ./template/parameters.json`

- az deployment group create --resource-group ResourceGroupName --template-file TemplateLocation

  --parameters ParametersLocation

  - e.g. Resource Group: David

    - az deployment group create --resource-group David --template-file ./template/template.json

      --parameters ./template/parameters.json
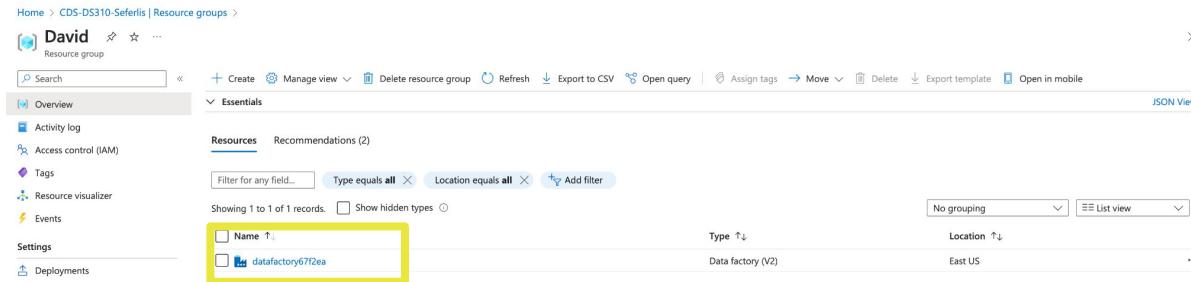
HOMEWORK 1B

Extract & Load

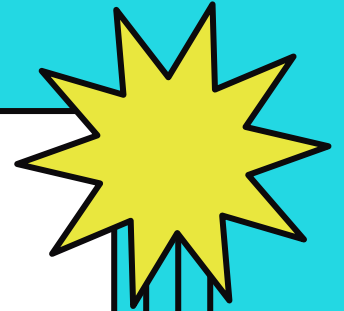| 01	02	03
Data Factory	Extract & Load	Unzip

1. DATA FACTORY

- **Large kitchen** with recipes coming from all over the world

  - You're one of the **Chefs**

    - **Gathering Ingredients** (Data Ingestion)

    - **Preparing Ingredients** (Data Transformation)

    - **Cooking** (Data Computation)

    - **Serving the Dish** (Data Publication)
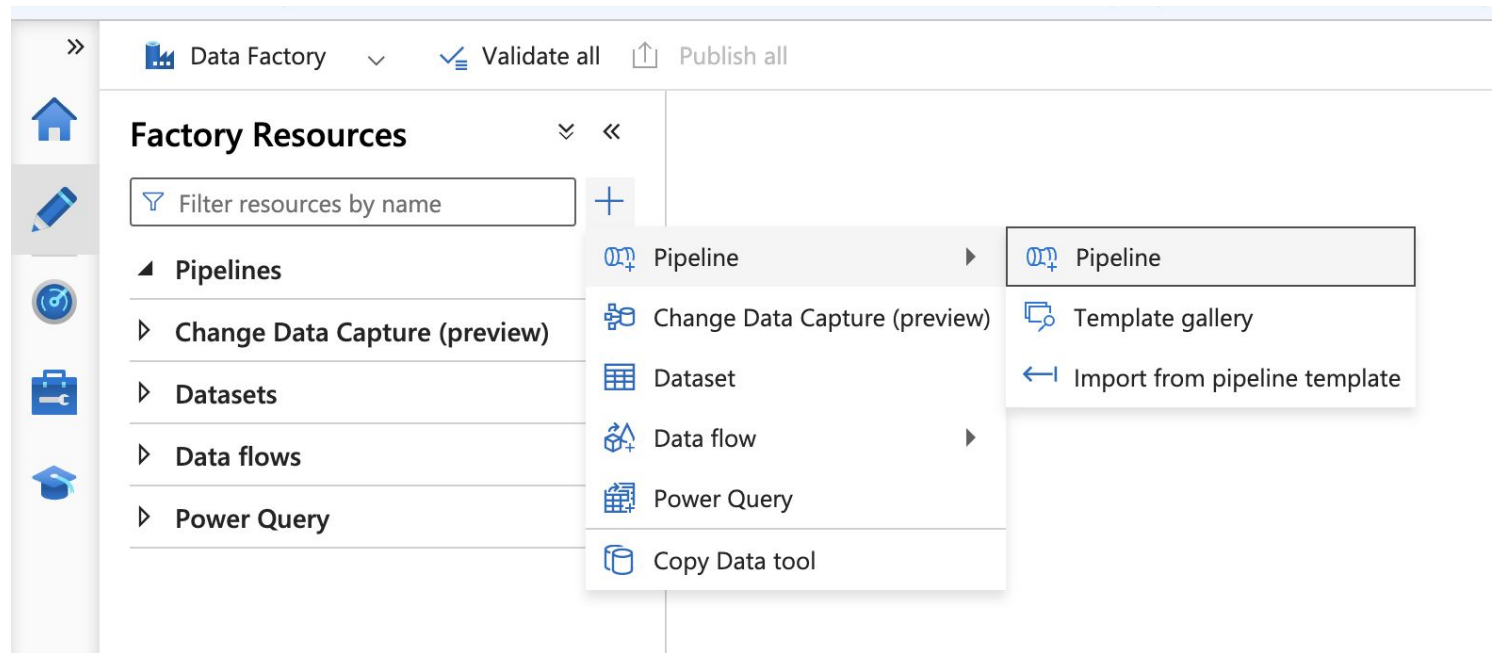
1. DATA FACTORY

Watch this video at home!

https://www.microsoft.com/en-us/videoplayer/embed/RE4Mc3u
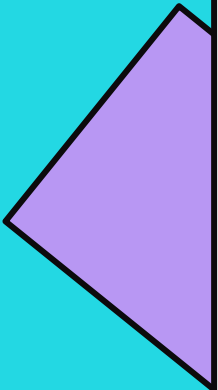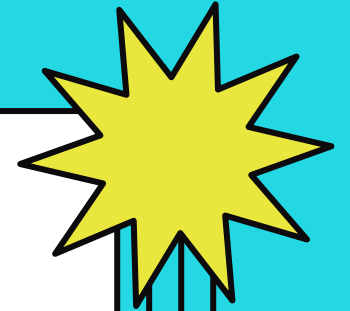
HOMEWORK 1B

Creating a Pipeline

HOMEWORK 1B

Creating Activities

Activities ⌄ «

Filter resources by name +

◢ Pipelines 1

⬤ pipeline1

▷ Change Data Capture (preview) 0

▷ Datasets 0

▷ Data flows 0

▷ Power Query 0

🔍 Search activities

⌄ Move and transform

▦ Copy data

🔀 Data flow

▷ Synapse

▷ Azure Data Explorer

▷ Azure Function

▷ Batch Service

▷ Databricks

▷ Data Lake Analytics

▷ General

▷ HDInsight

▷ Iteration & conditionals

▷ Machine Learning

▷ Power Query

01
COPY DATA 1

02
COPY DATA 2
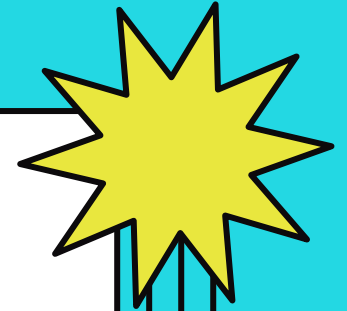
Pipeline

Extract & Load

Unzip

2. EXTRACT & LOAD

- Set up **Copy Data** activity to **Extract data** from

  - 'https://static.nhtsa.gov/odi/ffdd/cmpl/FLAT_CMPL.zip'

  - Hint: Zip file is a **binary** file!

1. Extract it from **HTTP** Linked Service

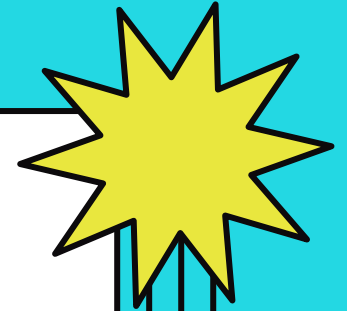2. Load it to a container in your **Azure Blob Storage**

HTTP

Azure Blob Storage

3. UNZIP

- **Unzip the file** in the Azure Blob Storage

- Hint:

  - Change the compression type

  - Change the .zip extension to…

  - Check Azure Storage Account Container for Output

    - Size of the txt file should be larger than that of the zip file

Azure Blob Storage

*RUNNING A PIPELINE

- **Publish & Add Trigger** >> Cheaper than 'Debug' Action

- Monitor pipeline runs through the **'monitor'** section

  - Go to details (eyeglasses icon) for details on data read & written)

BEST OF LUCK!