Decode HR

Power BI Dashboard User Guide

Contents

[1 Introduction 4](#_Toc40258855)

[2 Installation requirements 5](#_Toc40258856)

[2.1 Copy the dashboard from Google Drive to laptop 5](#_Toc40258857)

[2.2 Windows Installation 5](#_Toc40258858)

[3 Folders and files Information 6](#_Toc40258859)

[3.1 Windows Task Scheduling for running the excel download python code 7](#_Toc40258860)

[3.2 Files and Folders Summary 12](#_Toc40258861)

[3.3 Excel dataset download 12](#_Toc40258862)

[3.4 Configuration text input entry 13](#_Toc40258863)

[3.4.1 GoogleRSS input editing 13](#_Toc40258864)

[3.4.2 Twitter input editing 14](#_Toc40258865)

[4 Using Microsoft Power BI Dashboard: 16](#_Toc40258866)

Table of Figures

[Figure 3‑1 : Dashboard Folder and File structure 6](#_Toc40258914)

[Figure 3‑2 : GoogleNews Folder and File Structure 6](#_Toc40258915)

[Figure 3‑3 : Twitter Folder and File structure 6](#_Toc40258916)

[Figure 3‑4 : GoogleNews Data Folder and File Structure 7](#_Toc40258917)

[Figure 3‑5 : Twitter Data Folder and File Structure 7](#_Toc40258918)

[Figure 3‑6 : Dashboard Input Folder and File Structure 7](#_Toc40258919)

[Figure 3‑7 Running Windows task scheduler 8](#_Toc40258920)

[Figure 3‑8 : Create task for Windows task scheduler 8](#_Toc40258921)

[Figure 3‑9 : General tab - Task Name 9](#_Toc40258922)

[Figure 3‑10 : Triggers tab – Schedule time to run the task 9](#_Toc40258923)

[Figure 3‑11 : Actions tab – Select Batch file to run 9](#_Toc40258924)

[Figure 3‑12 : Conditions tab – Enable the network connections check box 10](#_Toc40258925)

[Figure 3‑13 : Settings tab 10](#_Toc40258926)

[Figure 3‑14 : Public HTML download folder 11](#_Toc40258927)

[Figure 3‑15 : Create User and Password for folder protection 11](#_Toc40258928)

[Figure 3‑16 : Google\_search\_Category.txt input file 13](#_Toc40258929)

[Figure 3‑17 : stopwords.txt 13](#_Toc40258930)

[Figure 3‑18 : spacy\_model.txt 14](#_Toc40258931)

[Figure 3‑19 : Twitter by Topic List.txt 14](#_Toc40258932)

[Figure 3‑20 : Twitter\_token.txt 15](#_Toc40258933)

[Figure 4‑1 : GoogleNews Power BI Files 16](#_Toc40258934)

[Figure 4‑2 : Power BI First Startup - Refresh datasets 16](#_Toc40258935)

[Figure 4‑3 : Change dataset locations 17](#_Toc40258936)

[Figure 4‑4 : Keyword Selection display on WordCloud 18](#_Toc40258937)

[Figure 4‑5 : WordCloud word selections 18](#_Toc40258938)

[Figure 4‑6 : Published Dates selections 19](#_Toc40258939)

[Figure 4‑7 : Articles Windows 19](#_Toc40258940)

[Figure 4‑8 : Topic Model tab 20](#_Toc40258941)

[Figure 4‑9 : 2 Words WordCloud Tab 20](#_Toc40258942)

[Figure 4‑10 : Figure 4‑11 : 3 Words WordCloud Tab 21](#_Toc40258943)

# Introduction

This document is for the documentation user guide for User to know how to setup and prepare for the Microsoft Power BI usage in Microsoft Windows 10 platform.

# Installation requirements

This section described the required software installation to support the operation.

## Copy the dashboard from Google Drive to laptop

User can directly copy from DCHR google drive folder named “dcHR.tech phase 3/dashboard-Copy\_to\_laptop” to local laptop C drive root directory. Rename the “dashboard-Copy\_to\_laptop” to “dashboard”. The total size of these folders is about 70MB

All the folders and required files are already created. The only action user would be required is to install the python and Microsoft Power BI. And add in the windows task schedule as describe in section 3.1

## Windows Installation

The software installation executable files are store in the dashboard/installation folder. Double click on the respective executable and follow the instructions.

1. Install python - python-3.8.2-amd64.exe (Under Installation folder)
   1. This is to install python version 3.8.2 in Windows 10 environment (Need to select the “Add to path” option checkbox during installation)
   2. This is to enable to run python codes to excel data file downloading from Support Board server
   3. The Python installation file requires about 25 Mb of disk space. When installed, Python requires about an additional 90 Mb of disk space. RAM: 1GB or more memory. Processors: Intel Atom® processor or Intel® Core™ i3 processor.
2. Install Microsoft Power BI - PBIDesktopSetup\_x64.exe (Under Installation folder)
   1. This is to install Microsoft Power BI software to able to run the dashboard application
   2. RAM- Minimum: 1 GB and Recommendation: 4 GB or more. Processor speed-Minimum x64 1.4GHz Recommended 2GHz or faster. Hard Disk- At least 1 GB spac

# Folders and files Information

The main folders information as follow: (As shown in Figure 5‑1). Both Windows and Linux platform share these same folders structure.

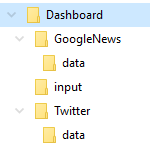
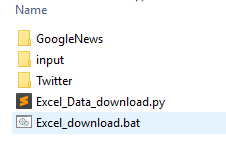
 

Figure 3‑1 : Dashboard Folder and File structure

1. Installations Folder
   1. Python and Microsoft Power BI Installations files
2. GoogleNews Folder for Power BI files
   1. GoogleNews\_by\_Category.pbix
   2. GoogleNews\_by\_Keyword.pbix
   3. GoogleNews\_by\_Product.pbix

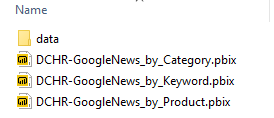


Figure 3‑2 : GoogleNews Folder and File Structure

1. Twitter Folder Power BI files
   1. Twttier\_by\_topic.pbix
   2. Twttier\_influencer.pbix

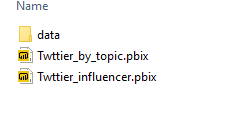


Figure 3‑3 : Twitter Folder and File structure

1. GoogleNews Folder create data folder to store the excel datasets
   1. DCHR\_Google\_Category.xlsx
   2. DCHR\_Google\_Keyword.xlsx
   3. DCHR\_Google\_Product.xlsx

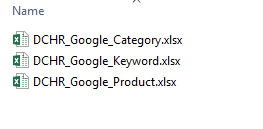


Figure 3‑4 : GoogleNews Data Folder and File Structure

1. Twitter Folder create data folder to store the excel datasets
   1. DCHR\_Twitters\_by\_topics.xlsx
   2. DCHR\_Influencer\_tweets.xlsx

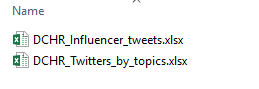


Figure 3‑5 : Twitter Data Folder and File Structure

1. Input folder to store the user editable text files for web server address and local address configuration and CPanel Username & Password for the access of downloading
   1. Cpanel\_pwd.txt
   2. src\_dest.txt

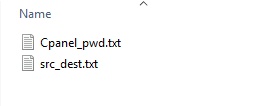


Figure 3‑6 : Dashboard Input Folder and File Structure

## Windows Task Scheduling for running the excel download python code

1. Run the windows task scheduler as shown in Figure 3‑7
2. Create new task as shown in Figure 3‑8
3. Enter the desire task name in the General tab as shown in Figure 3‑9
4. Configure the desire schedule in the Trigger tab to run the task as shown in Figure 3‑10
5. Select the desire batch file to run as shown in Figure 3‑11
6. Set the desire conditions in the condition tab as shown in Figure 3‑12
7. Set the desire settings in the Steeings tab as shown in Figure 3‑13

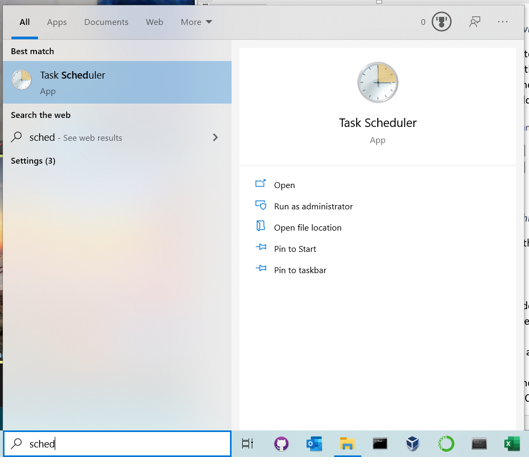


Figure 3‑7 Running Windows task scheduler

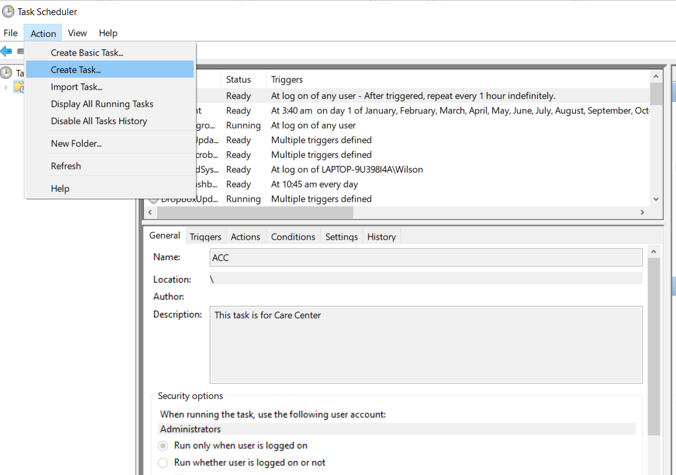


Figure 3‑8 : Create task for Windows task scheduler

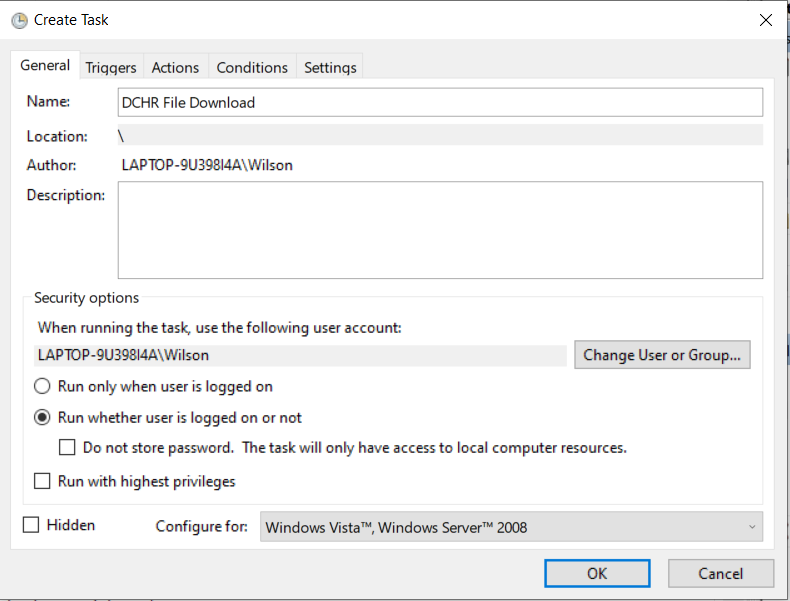


Figure 3‑9 : General tab - Task Name

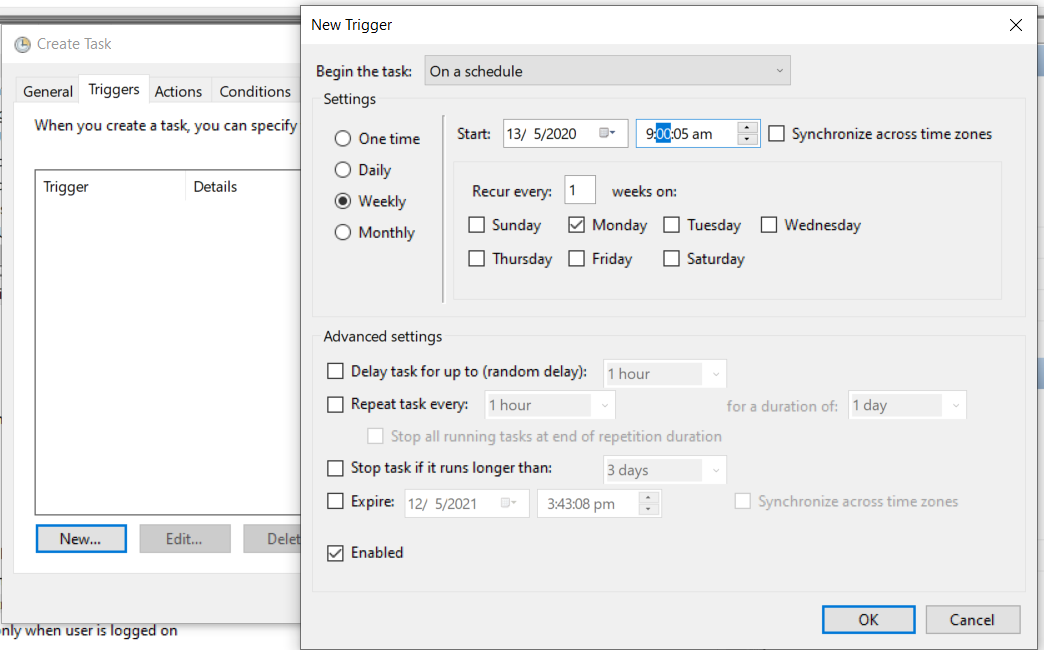


Figure 3‑10 : Triggers tab – Schedule time to run the task

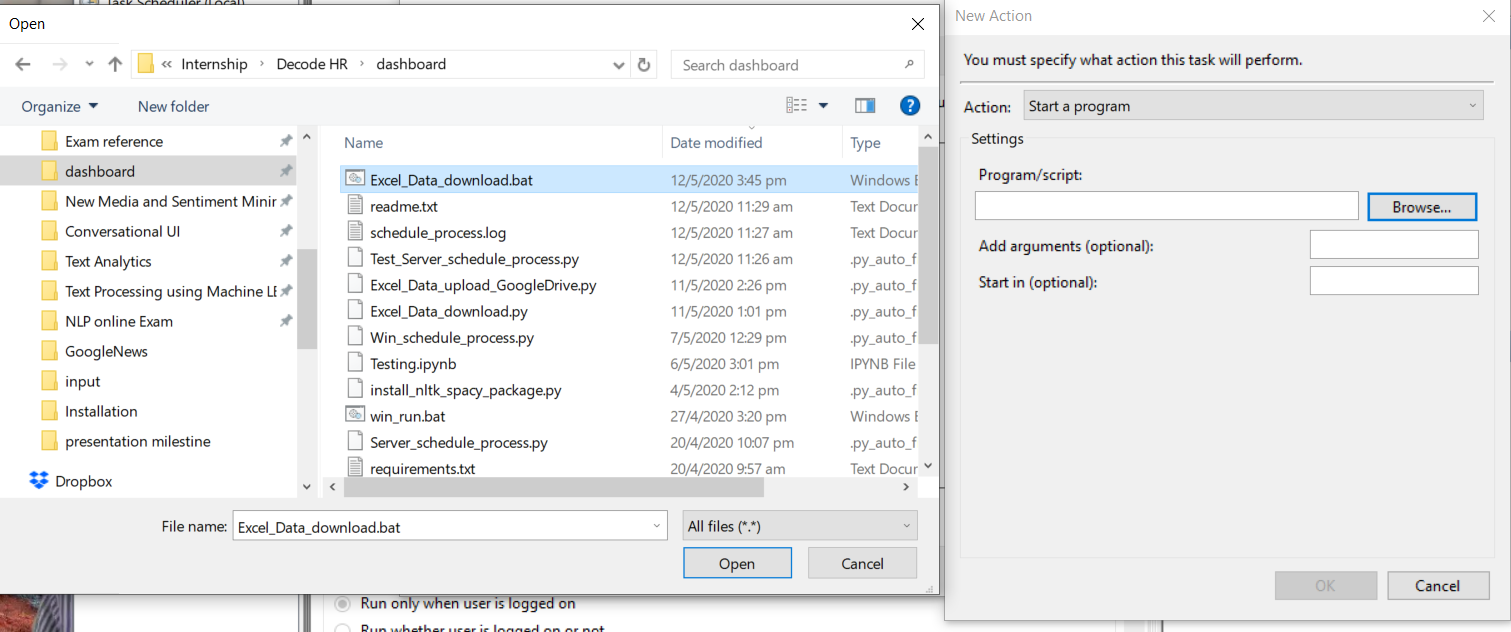


Figure 3‑11 : Actions tab – Select Batch file to run

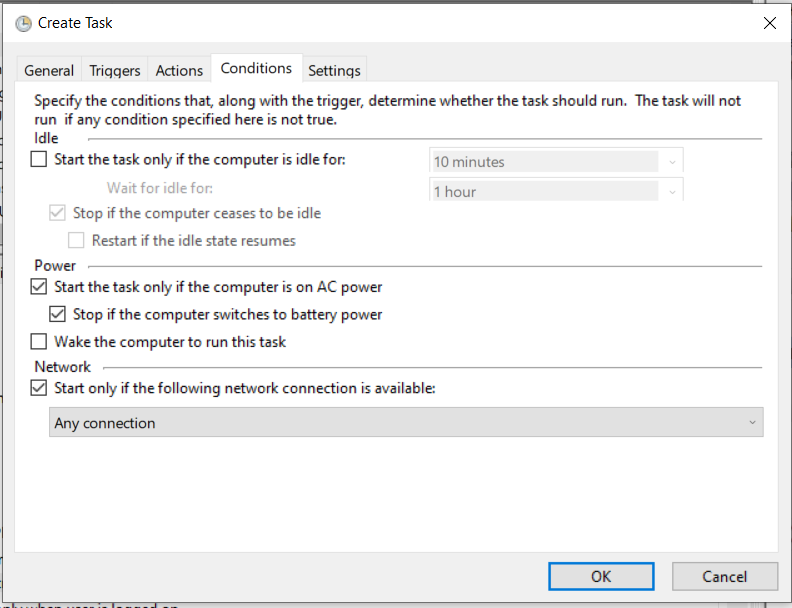


Figure 3‑12 : Conditions tab – Enable the network connections check box

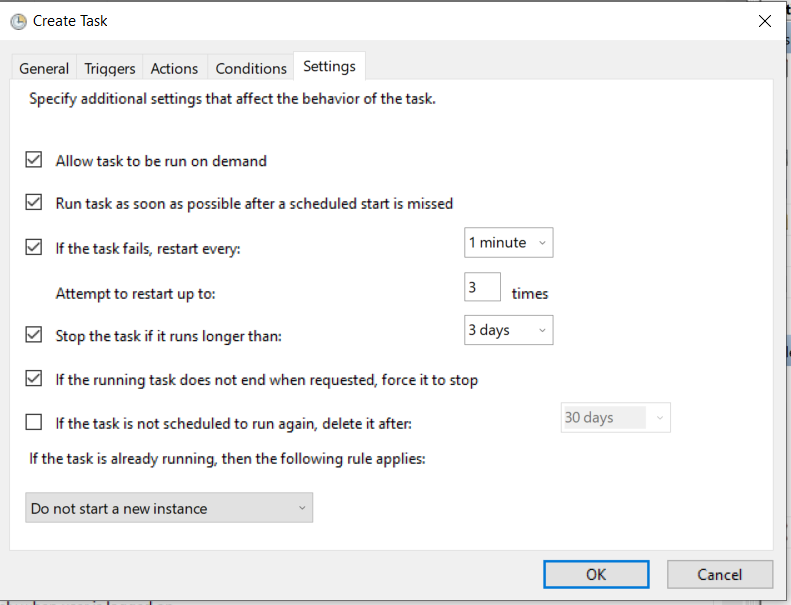


Figure 3‑13 : Settings tab

1. This is for changing of Public HTML folder Username and password
   1. Click on the File Manager in the CPanel main page in the server
   2. Go to public\_html folder
   3. Right Click on the DCHR folder and select “Password Protect” in the popup menu as shown in Figure 3‑7
   4. Key in the desire User name and password and click save
   5. The next step is to update the Cpanel.txt file for the user name and password in the input folder

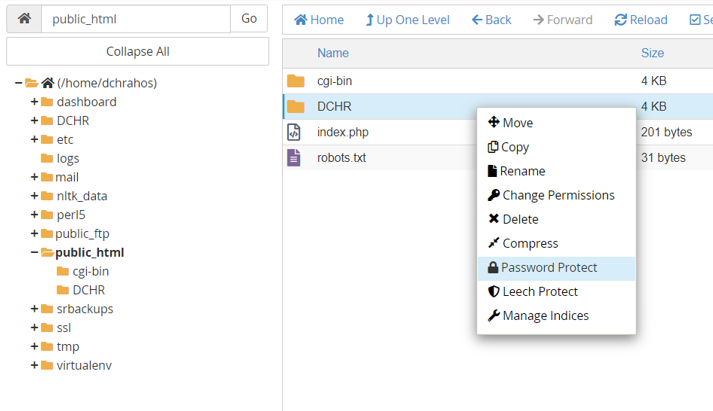


Figure 3‑14 : Public HTML download folder

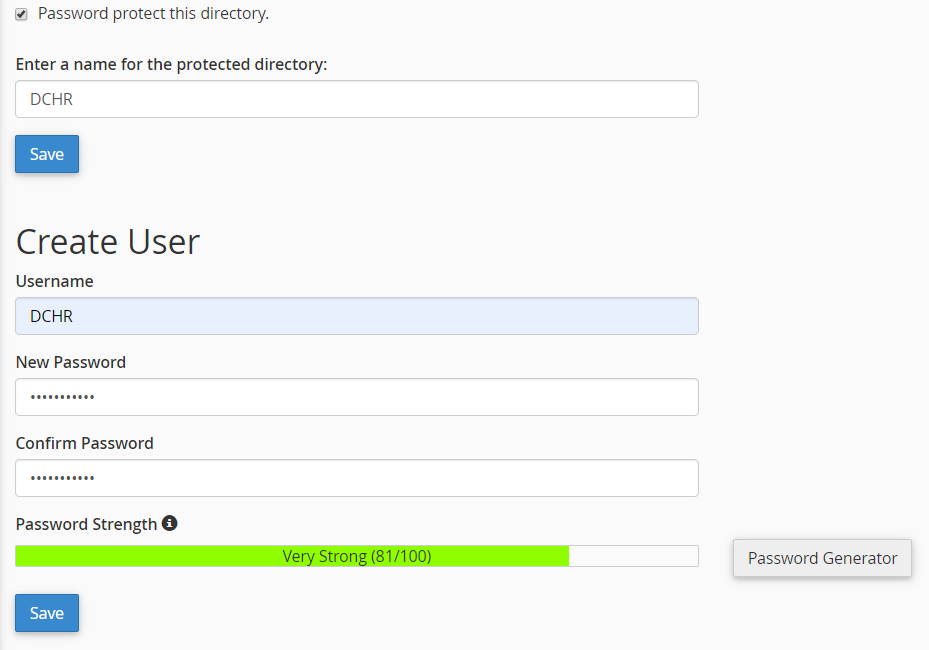


Figure 3‑15 : Create User and Password for folder protection

## Files and Folders Summary

User are required to create the following folders:

1. Dashboard
2. Under dashboard folder create these folder:
   1. GoogleNews
   2. twitter
3. Under GoogleNews folder upload these Power BI files:
   1. GoogleNews\_by\_Category.pbix
   2. GoogleNews\_by\_Keyword.pbix
   3. GoogleNews\_by\_Product.pbix
4. Under GoogleNews folder upload
5. These Under GoogleNews folder creates:
   1. Data
6. Under GoogleNews/data folder, user will have to upload the following excel files from server to this folder:
   1. DCHR\_Google\_Category.xlsx
   2. DCHR\_Google\_Keyword.xlsx
   3. DCHR\_Google\_Product.xlsx
7. Under GoogleNews folder upload these Power BI files:
   1. Twttier\_by\_topic.pbix
   2. Twttier\_influencer.pbix
8. Under twitter folder creates:
   1. Data
9. Under twitter/data folder, user will have to upload the following excel files from server to this folder:
   1. DCHR\_Twitters\_by\_topics.xlsx
   2. DCHR\_Influencer\_tweets.xlsx
10. Under Dashboard folder creates:
    1. Input
11. Under dashboard/input folder upload the files:
    1. Cpanel\_pwd.txt – For authentication to Support Board server
    2. src\_dest.txt – for source and destination address for the excel file downloading
12. Then go to section 8 to learn how to use Power BI

## Excel dataset download

User are expected to have install the python. Run the python-3.8.2-amd64.exe and using all default settings. The run the following windows batch file:

* Excel\_download.bat

This will download all the twitter and GoogleRss News excel file from support board server to local respective data folder as mentioned section 5.1.

Take note that editing is require in the batch file to reflect the local folder path to dashboard folder.

## Configuration text input entry

Users are expected to have install the python. Run the python-3.8.2-amd64.exe and using all default settings. The run the following

### GoogleRSS input editing

The Category, Keywords and Products editing text file will be in the dashboard/GoogleNews/input folder. For new keywords to scrap, need to input into this text file line by line. Example as shown in Figure 3‑7 for Category and it applies the same for Keyword and Product.

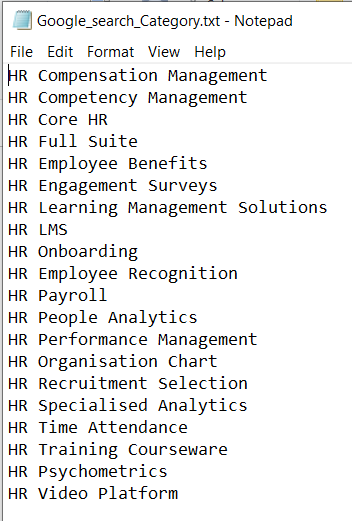


Figure 3‑16 : Google\_search\_Category.txt input file

#### Stopwords editing

This stopword input text file is meant for all Category, Keyword and Product preprocessing of removal of undesired words to appear in the wordcloud. For new stopwords to remove, need to input into this text file line by line and a word per line. As the stopwords only words to remove one word at a time not phrases.

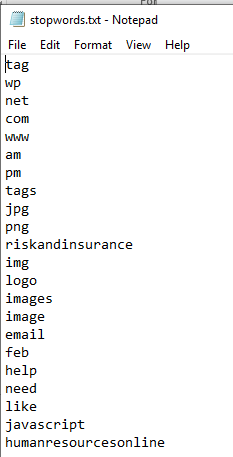


Figure 3‑17 : stopwords.txt

#### Spacy dictionary size entry

This input is to select the python vocabulary database size to be used for processing the text data scrap from GoogleRSS site. There are 3 choices namely :

1. en\_core\_web\_sm – small size about 10MB
2. en\_core\_web\_md – medium size about 100MB
3. en\_core\_web\_lg – large size about 800MB

Default use in the program will be en\_core\_web\_sm (Small size) due to server capacity. If server capacity is increase, the other 2 size can be used. Only enter one option for this text file.

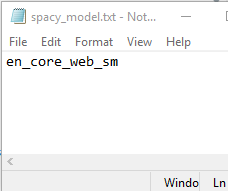


Figure 3‑18 : spacy\_model.txt

### Twitter input editing

The topic and Influencer editing text file will be in the dashboard/twitter/input folder. For new keywords to scrap, need to input into this text file line by line. Example as shown in Figure 3‑7 for topic and it applies the same for Influencer.

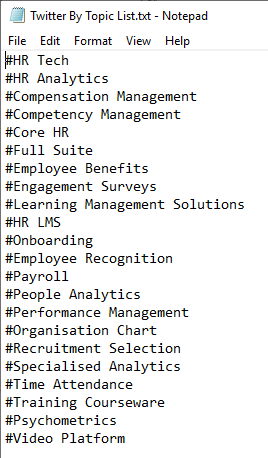


Figure 3‑19 : Twitter by Topic List.txt

#### Twitter token editing

Twitter scrapping require the user login credential and aaccess tokens which can be applied from Twitter developer web site and the program will read in these tokesn from this twitter\_token.txt to access and scrape via the Twitter API.



Figure 3‑20 : Twitter\_token.txt

# Using Microsoft Power BI Dashboard:

1. Run or double click the “DCHR-GoogleNews\_by\_Category.pbix” file to open up the Microsoft Power BI application

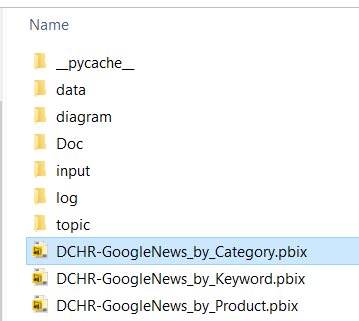


Figure 4‑1 : GoogleNews Power BI Files

1. First Page on the Power BI Dashboard “Google News Article” shows the overall dashboard of the Google News for all the HR Category. Click on the “Refresh” button on the Menu for updating the data from the latest updates.

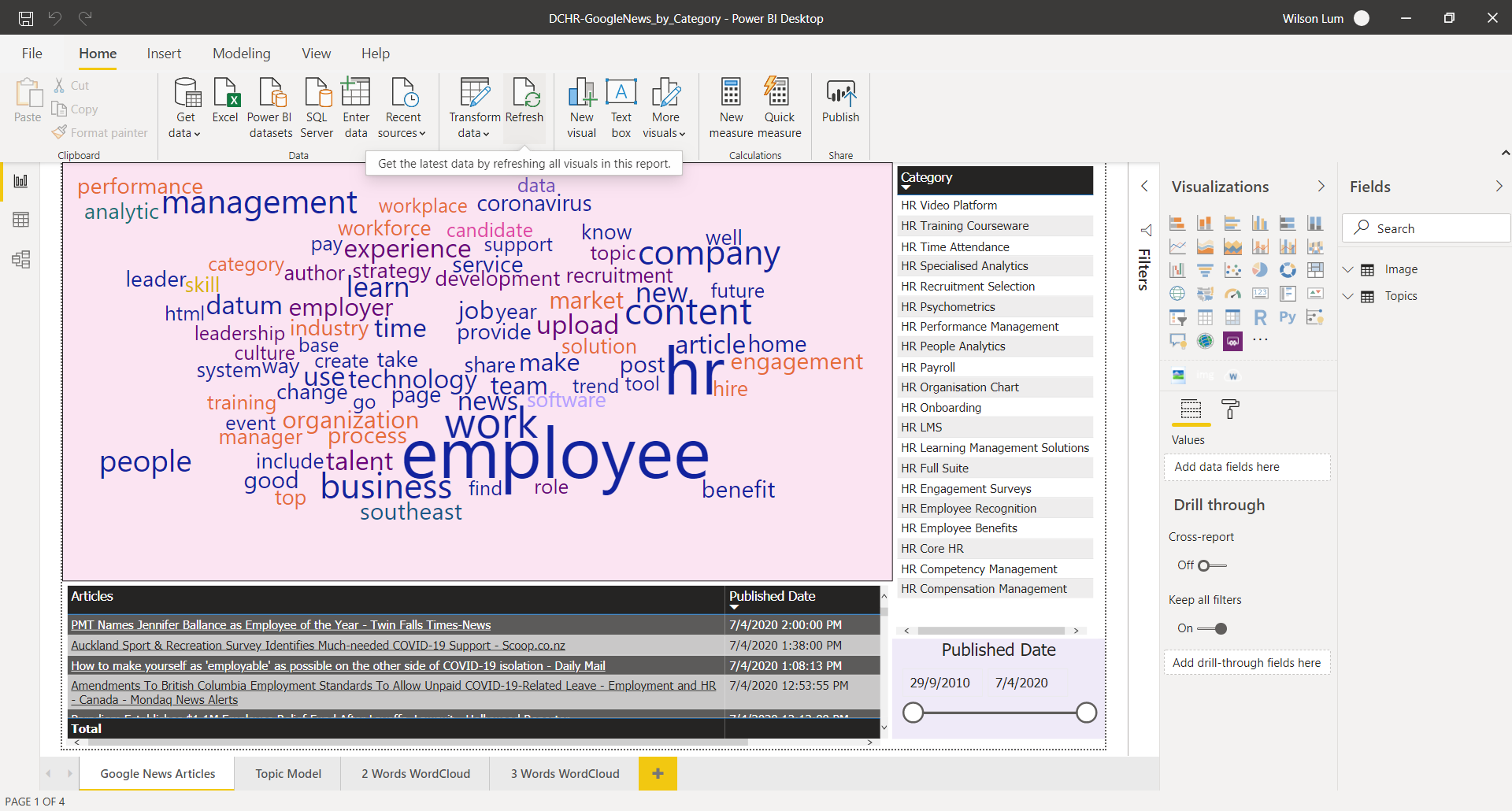


Figure 4‑2 : Power BI First Startup - Refresh datasets

1. IF for there is a change in the dataset folder, go to the “Transform Data” menu and select “Data Source Setting” and change to the desire dataset location. By default, the excel data file is located in the dashboard/GoogleNews/data.

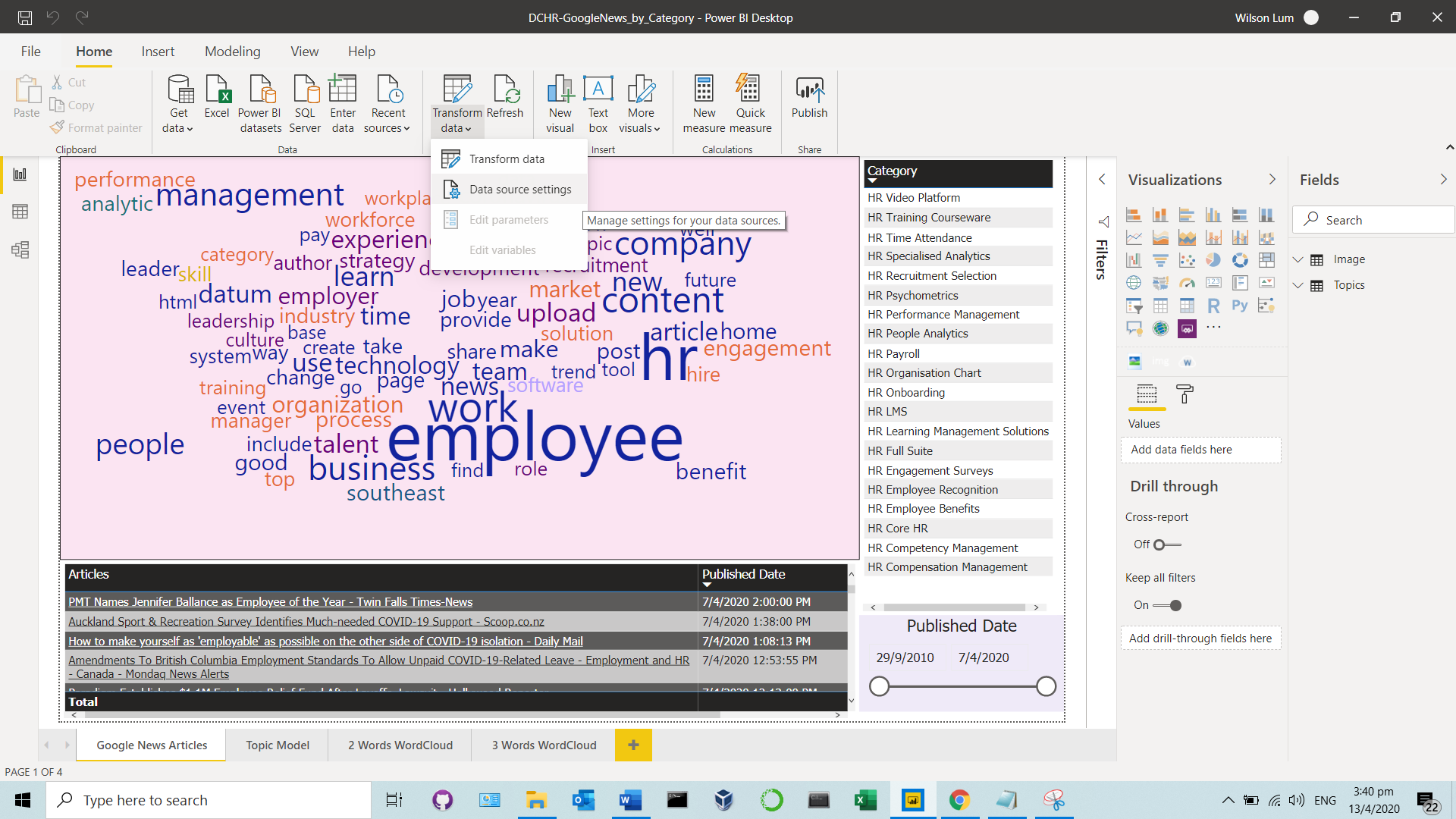


Figure 4‑3 : Change dataset locations

1. Top Right Window - “Category” Selection
   1. In this window, it allows one to selected the desire category. Once selected, the action will trigger the rest of the two windows, “WordCloud” and “Articles & Published Date” to updates information with reference to the selected category
   2. When one category is selected, the rest will be greyed out and the “WordCloud” and “Articles & Published Date” windows will refresh and updated
   3. To return to original state, click on the selected category word to deselect.



Figure 4‑4 : Keyword Selection display on WordCloud

1. Top Left Window – “WordCloud” Display
   1. Point the mouse cursor to the desire word in the wordcloud and it will display the number of times it appear in the datasets (Depending on the selected Category)
   2. Click on the word in the wordcloud and the action will trigger the rest of the two windows, “Category” and “Articles & Published Date” to updates information with reference to the selected word.



Figure 4‑5 : WordCloud word selections

1. Bottom Left Window “Published Date” Selection
   1. The “Published date” window on the bottom left window is for the dates range selection with respect to the datasets.
   2. Use the mouse cursor to move the bubble left and right to adjust the desire dates,

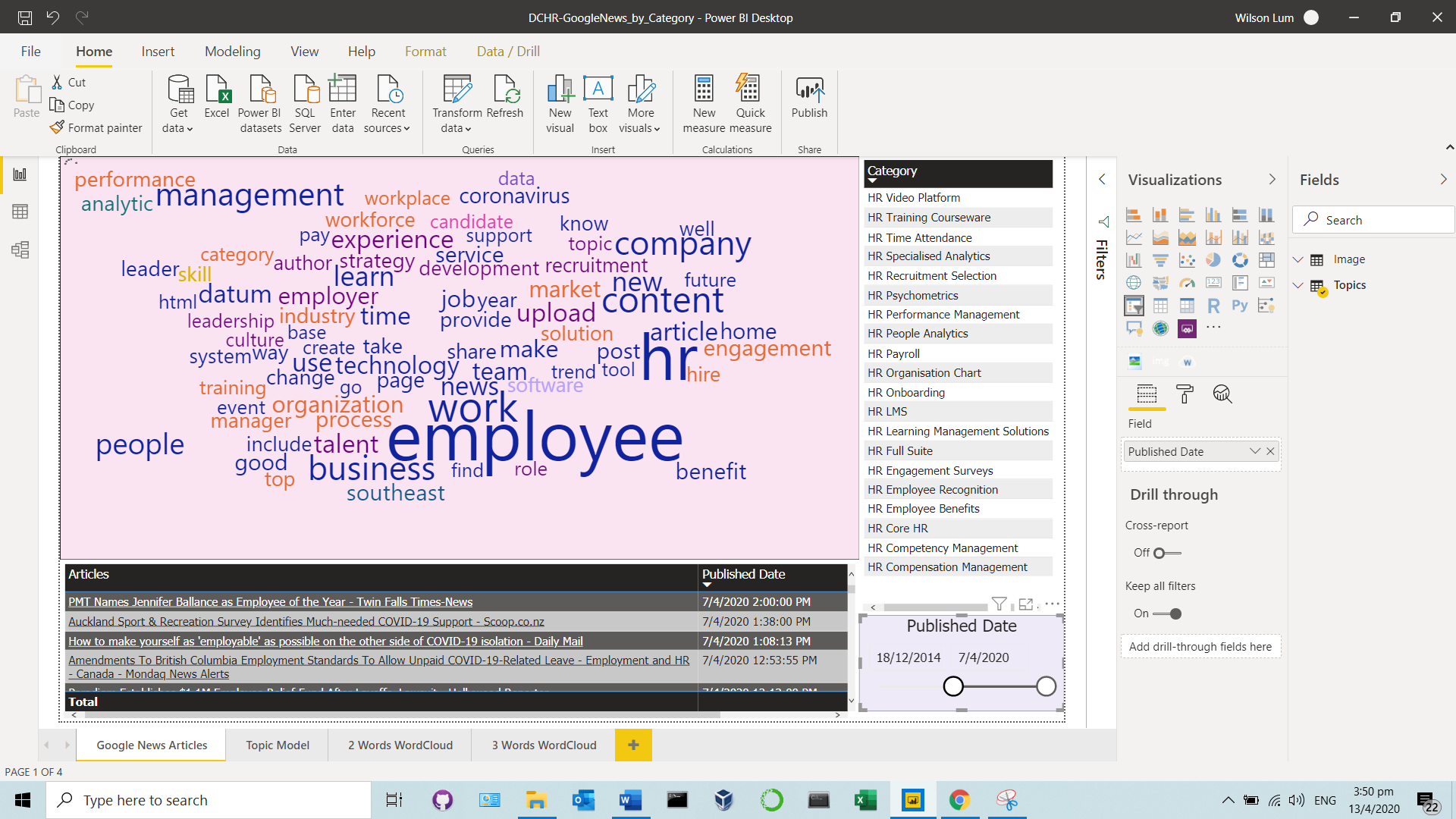


Figure 4‑6 : Published Dates selections

1. Bottom Left Window – “Articles & Published Date” Selections
   1. Move the mouse to any of the interested text and it will display the link info
   2. Click on the underline text (Which is a hyperlink) and the browser will launch and visiting to the respective link.

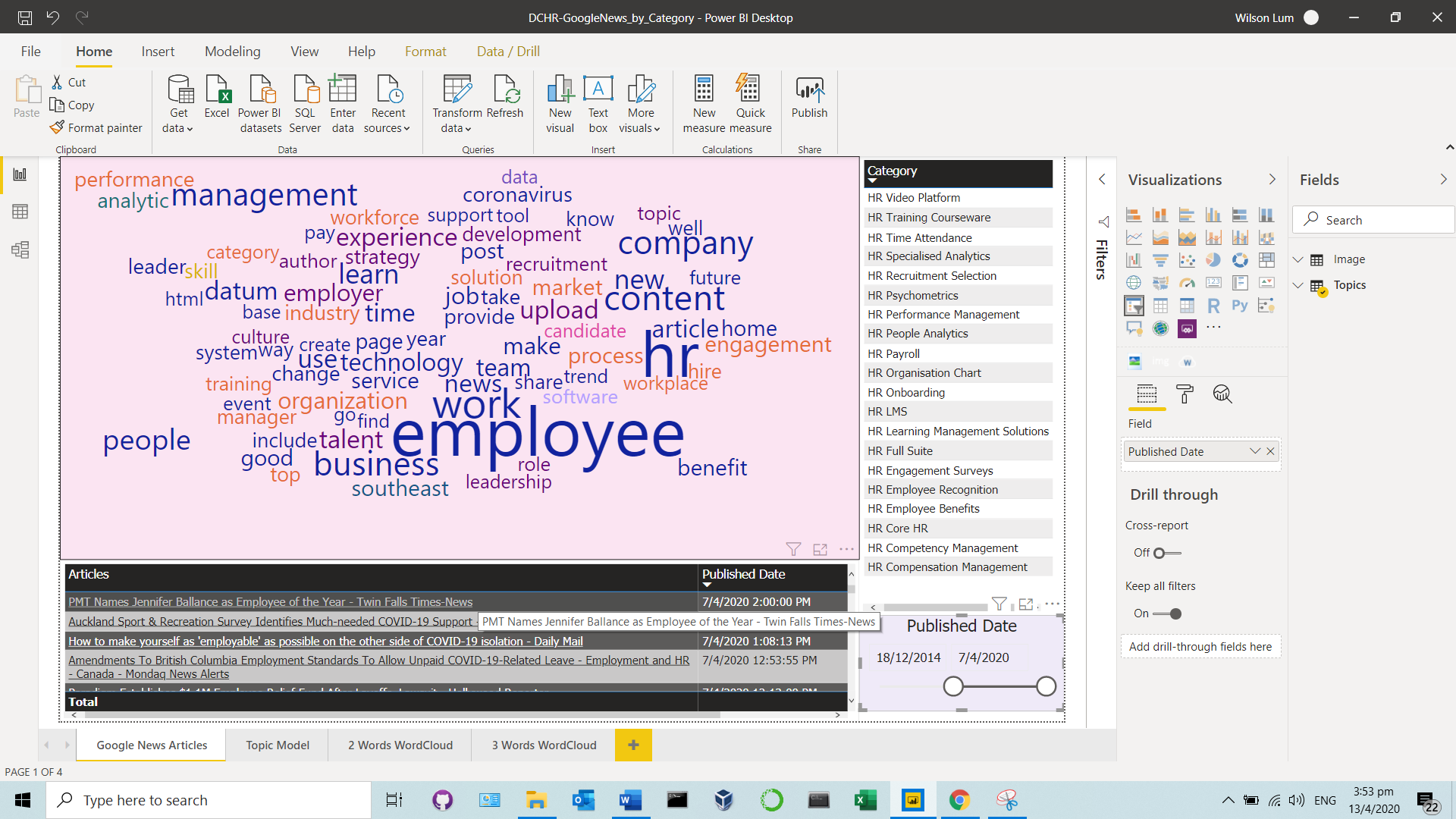


Figure 4‑7 : Articles Windows

1. Topic Model Tab
   1. This tab has major difference from the “Google News Articles” that it has 3 sub wordcloud and each sub wordcloud display a topic of its own.

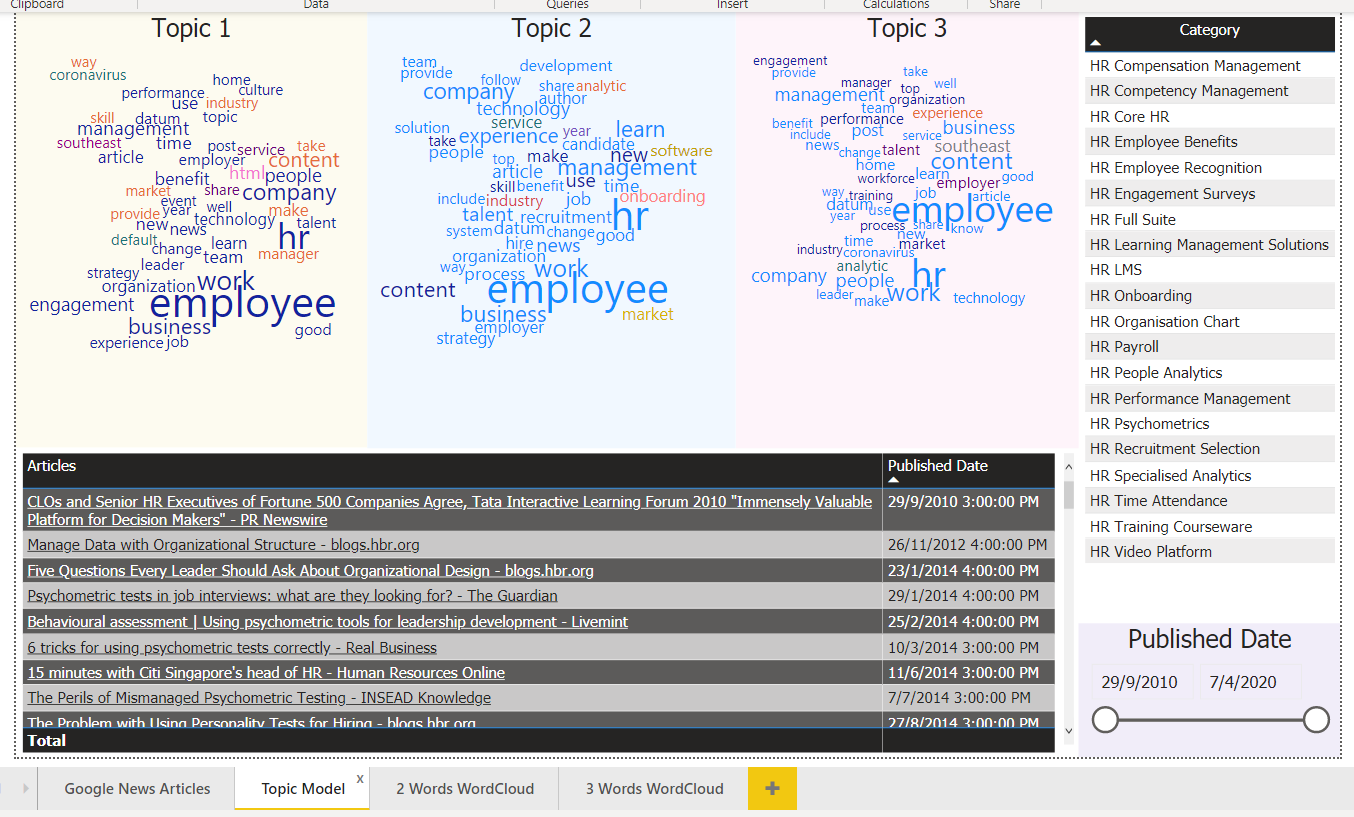


Figure 4‑8 : Topic Model tab

1. 2 Words Wordcloud Tab
   1. This tab display the bigram or 2 words wordcloud for each selected category

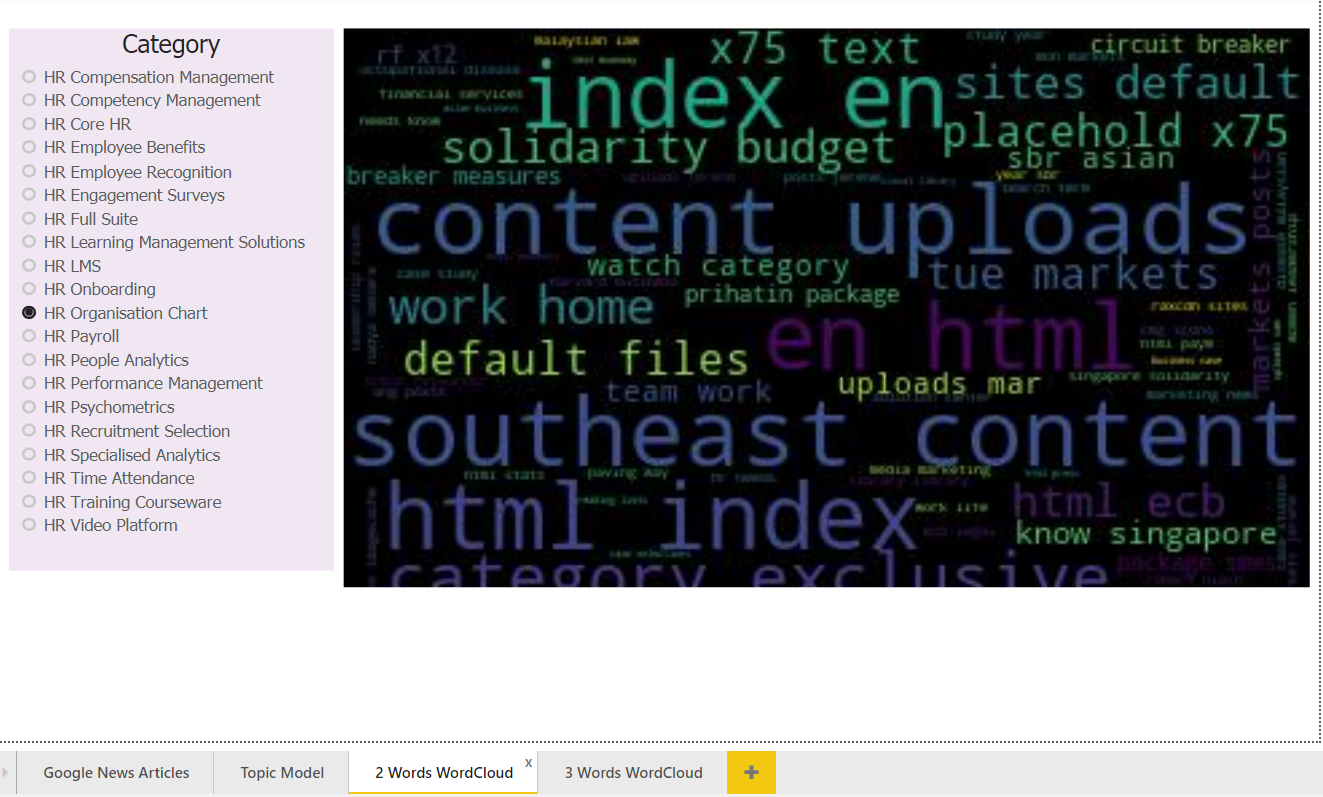


Figure 4‑9 : 2 Words WordCloud Tab

1. 3 Words Wordcloud Tab
   1. This tab display the trigram or 3 words wordcloud for each selected category



Figure 4‑10 : Figure 4‑11 : 3 Words WordCloud Tab