

How to automate Excel update using Jupyter Notebook integration within Cognos Analytics.

Janne Lind, IBM
June 5, 2020

This document describes how to use [Jupyter Notebooks](#) in IBM Cognos Analytics to create a new data set from an Excel source.

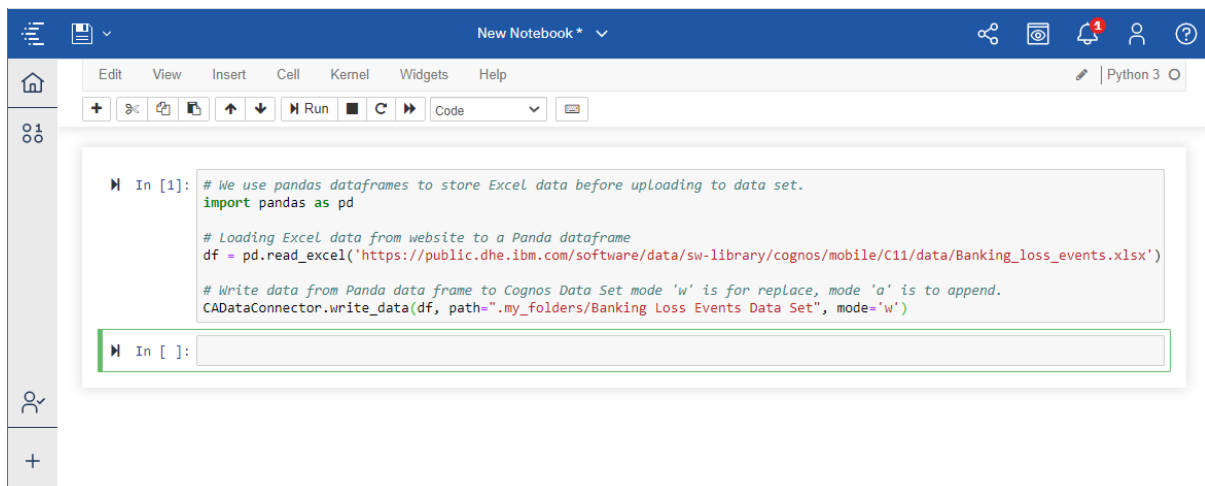
Step 1. Create Notebook.

Here is very simple Python example that loads an Excel file from a website. You can replace the *Banking_loss_events.xlsx* file URL reference with any Excel file that is in tabular format (see [Uploaded files](#)).

```
# We use pandas dataframes to store Excel data before uploading to data set.
import pandas as pd

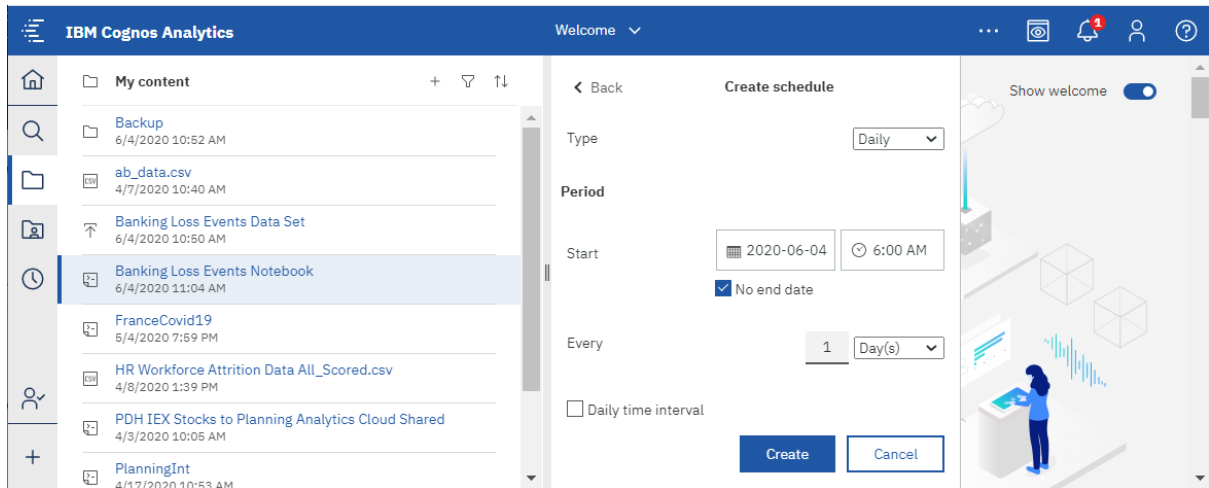
# Loading Excel data from website to a Panda dataframe
df = pd.read_excel('https://public.dhe.ibm.com/software/data/sw-library/cognos/mobile/C11/data/Banking_loss_events.xlsx')

# Write data from Panda data frame to Cognos Data Set mode 'w' is for replace, mode 'a' is to append.
CADataConnector.write_data(df, path=".my_folders/Banking Loss Events Data Set", mode='w')
```



Step 2. Create Schedule.

Create a schedule to run the Jupyter Notebook at your desired interval to update the data set. If the source Excel file is modified between the scheduled events, the data set will reflect these changes. For more information, see [Schedule Management](#).



Appendix A: Other File Types

The Python script can be changed to work with CSV and many other formats. You will just need to replace **pandas.read_excel** with **pandas.read_CSV** and update the file reference URL to point to the CSV file.

Here is a list of **pandas.read** options (some of which may require more changes to the code than simply replacing **pandas.read** function):

- pandas.read_clipboard
- pandas.read_csv
- pandas.read_excel
- pandas.read_feather
- pandas.read_fwf
- pandas.read_gbq
- pandas.read_hdf
- pandas.read_html
- pandas.read_json
- pandas.read_orc
- pandas.read_parquet
- pandas.read_pickle
- pandas.read_sas
- pandas.read_spss
- pandas.read_sql
- pandas.read_sql_query
- pandas.read_sql_table
- pandas.read_stata
- pandas.read_table