data-pipelining-with-coinbase

July 28, 2025

1 Data Pipelining With Coinbase

1.1 Python Imports

[1]:
```
# Standard Library
import datetime
import io
import os
import random
import sys
import warnings

from datetime import datetime, timedelta
from pathlib import Path

# Data Handling
import numpy as np
import pandas as pd

# Data Visualization
import matplotlib.dates as mdates
import matplotlib.pyplot as plt
import matplotlib.ticker as mtick
import seaborn as sns
from matplotlib.ticker import FormatStrFormatter, FuncFormatter, MultipleLocator

# Data Sources
import yfinance as yf

# Statistical Analysis
import statsmodels.api as sm

# Machine Learning
from sklearn.decomposition import PCA
from sklearn.preprocessing import StandardScaler

# Suppress warnings
warnings.filterwarnings("ignore")
```

1.2 Add Directories To Path

[2]:
```
# Add the source subdirectory to the system path to allow import config from␣
 ↪settings.py
current_directory = Path(os.getcwd())
website_base_directory = current_directory.parent.parent.parent
src_directory = website_base_directory / "src"
sys.path.append(str(src_directory)) if str(src_directory) not in sys.path else␣
 ↪None

# Import settings.py
from settings import config

# Add configured directories from config to path
SOURCE_DIR = config("SOURCE_DIR")
sys.path.append(str(Path(SOURCE_DIR))) if str(Path(SOURCE_DIR)) not in sys.path␣
 ↪else None

# Add other configured directories
BASE_DIR = config("BASE_DIR")
CONTENT_DIR = config("CONTENT_DIR")
POSTS_DIR = config("POSTS_DIR")
PAGES_DIR = config("PAGES_DIR")
PUBLIC_DIR = config("PUBLIC_DIR")
SOURCE_DIR = config("SOURCE_DIR")
DATA_DIR = config("DATA_DIR")
DATA_MANUAL_DIR = config("DATA_MANUAL_DIR")

# Print system path
for i, path in enumerate(sys.path):
    print(f"{i}: {path}")
```

```
0: /usr/lib/python313.zip
1: /usr/lib/python3.13
2: /usr/lib/python3.13/lib-dynload
3:
4: /home/jared/python-virtual-envs/general_313/lib/python3.13/site-packages
5: /home/jared/Cloud_Storage/Dropbox/Websites/jaredszajkowski.github.io/src
```

1.3 Track Index Dependencies

[3]:
```
# Create file to track markdown dependencies
dep_file = Path("index_dep.txt")
dep_file.write_text("")
```

[3]: 0

1.4 Python Functions

[4]:
```
from coinbase_fetch_available_products import coinbase_fetch_available_products
from coinbase_fetch_full_history import coinbase_fetch_full_history
from coinbase_fetch_historical_candles import coinbase_fetch_historical_candles
from coinbase_pull_data import coinbase_pull_data
from export_track_md_deps import export_track_md_deps
```

1.5 Function Usage

1.5.1 Coinbase Fetch Available Products

[5]:
```
df = coinbase_fetch_available_products(
    base_currency=None,
    quote_currency="USD",
    status="online",
)
```

[6]:
```
# Copy this <!-- INSERT_coinbase_fetch_available_products_HERE --> to␣
 ↪index_temp.md
export_track_md_deps(dep_file=dep_file,␣
 ↪md_filename="coinbase_fetch_available_products.md", content=df.
 ↪to_markdown(floatfmt=".5f"))
```

```
Exported and tracked: coinbase_fetch_available_products.md
```

1.5.2 Coinbase Fetch Historical Candles

[7]:
```
df = coinbase_fetch_historical_candles(
    product_id="BTC-USD",
    start=datetime(2025, 1, 1),
    end=datetime(2025, 1, 1),
    granularity=86_400,
)
```

[8]:
```
# Copy this <!-- INSERT_coinbase_fetch_historical_candles_HERE --> to␣
 ↪index_temp.md
export_track_md_deps(dep_file=dep_file,␣
 ↪md_filename="coinbase_fetch_historical_candles.md", content=df.
 ↪to_markdown(floatfmt=".5f"))
```

```
Exported and tracked: coinbase_fetch_historical_candles.md
```

1.5.3 Coinbase Fetch Full History

[9]:
```
df = coinbase_fetch_full_history(
    product_id="BTC-USD",
    start=datetime(2025, 1, 1),
    end=datetime(2025, 1, 31),
```

```
    granularity=86_400,
)
```

[10]:
```
# Copy this <!-- INSERT_coinbase_fetch_full_history_HERE --> to index_temp.md
export_track_md_deps(dep_file=dep_file,␣
 ↪md_filename="coinbase_fetch_full_history.md", content=df.
 ↪to_markdown(floatfmt=".5f"))
```

```
Exported and tracked: coinbase_fetch_full_history.md
```

1.5.4 Coinbase Pull Data

[11]:
```
# df = coinbase_pull_data(
#     base_directory=DATA_DIR,
#     source="Coinbase",
#     asset_class="Cryptocurrencies",
#     excel_export=False,
#     pickle_export=True,
#     output_confirmation=True,
#     base_currency="BTC",
#     quote_currency="USD",
#     granularity=60, # 60=minute, 3600=hourly, 86400=daily
#     status='online', # default status is 'online'
#     start_date=datetime(current_year, current_month - 1, 1), # default start␣
 ↪date
#     end_date=datetime.now() - timedelta(days=1), # updates data through 1 day␣
 ↪ago due to lag in data availability
# )
```