

ETF Data Product Manager 90 Second Pitch

Thanks for taking the time – just to give a quick background of me, I’m originally from New Jersey and studied undergrad at Penn State.

After graduating, I began my first role at Bloomberg L.P. as an Equity Research Data Analyst in TMT. This group is considered your traditional Sell Side Research Arm, covering stocks from a fundamentals approach. I managed a portfolio of 70 Media and Communications companies.

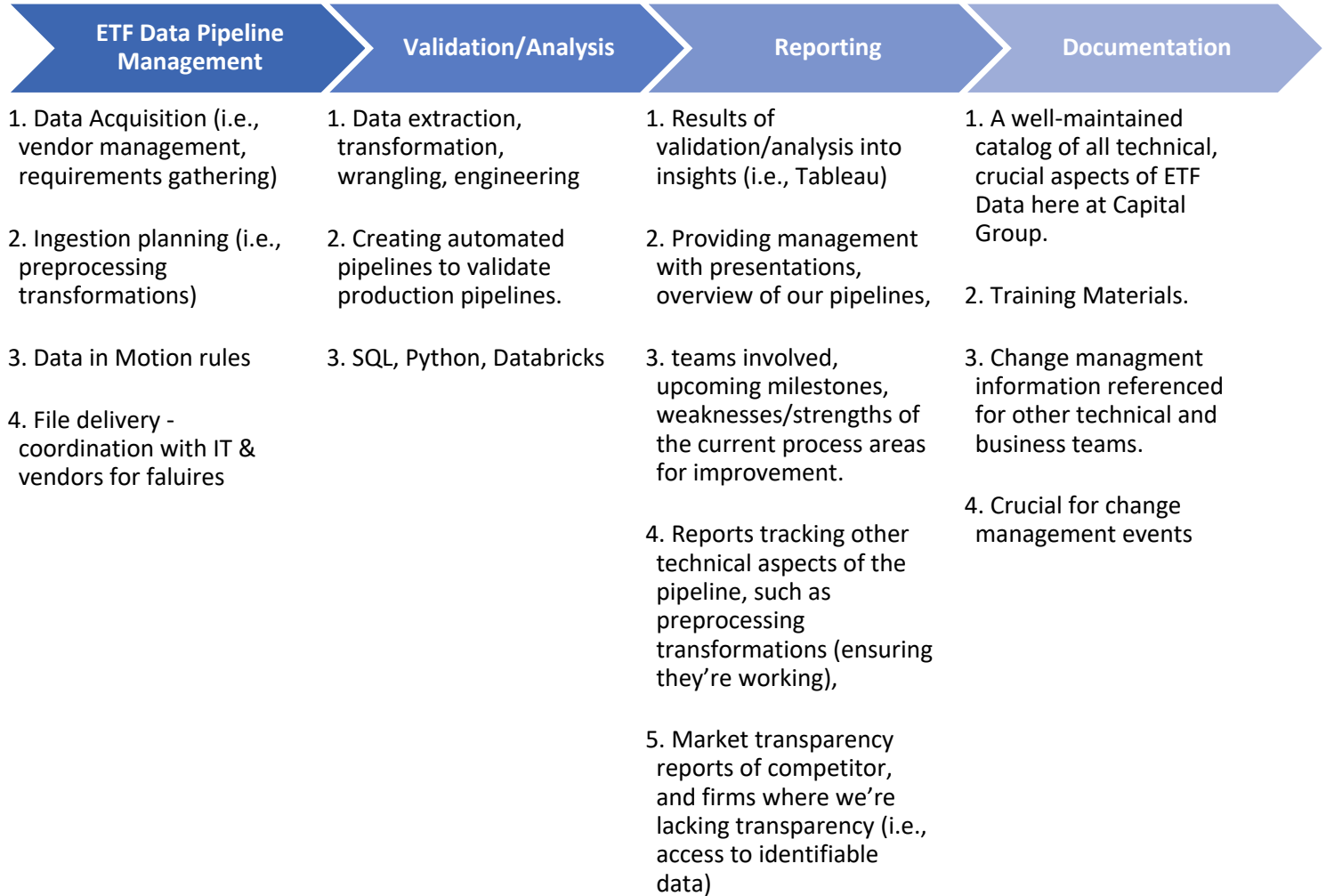
At Reorg Research in the capacity of a Data Analyst for M&A and Shareholder Activism, I played a pivotal role in developing a Machine Learning Model that scoured SEC filings for potential shareholder activist events. I championed the launch of our Pre-Deal-M&A Database, spotlighting 200+ companies under strategic review. Additionally, I refined our model's outputs, tracking the significant movements of 160+ targeted companies.

At Third Bridge Group as a Sector Analyst in Consumer Staples, I led over 125 high-level discussions with C-suite executives, diving deep into strategic positioning and consumer behavior. My proactive analysis expanded our content coverage across sectors, including Consumer Services and Agriculture, adding 142 equity names.

At Capital Group, as a Data Product Manager, I've elevated the landscape of our ETF data and analytics, enhancing the visibility and value of our \$8 billion product suite. My role involves bridging the gap between various domains—IT, Sales Compensation, Data Science, Engineering, Consulting, & Business Reporting—by creating synergies that maximize our data's potential.

Started as Data Product Analyst within Insights & Analytics > Promoted to Manager in Enterprise Data Office.

My Core Activities



Product - ETF Data & Pipeline

1. Data Acquisition

- a. Use case
- b. Requirements gathering
- c. Vendor outreach/communication
- d. Data quality checking (2 months)

2. Ingestion Planning

- a. ETL Pipeline
- b. 7-8 different systems
- c. 6-7 different teams
- d. 2 pipelines run simultaneously for Sales and one for Assets

3. Production – Data file deliver

- a. Land as MFT – Security platform
- b. Moved to DataLake – we call Caspian with is on AWS S3 platform
- c. Sales transactions processed through 4 additional systems
- d. Assets is sent to ADLS, transformed, the published to Data warehouse
- e. Ensure all raw untransformed data is delivered to ADLS

4. Validation

- a. Did our systems operate as intended?
- b. Are we missing data?
- c. Did the vendor meet SLA?
- d. Did we duplicate any data along the way?
- e. How long did it take from start to finish?
- f. Did we perform the right transformations (i.e., block seeding account numbers)?

List of Responsibilities

1. ETF Data Pipeline Management

- a. Data Acquisition (i.e., vendor management, requirements gathering)
- b. Ingestion planning (i.e., preprocessing transformations)
- c. Data in Motion

2. Validation/Analysis

- a. Data extraction, transformation, wrangling, engineering
- b. Creating automated pipelines to validate production pipelines.
- c. SQL, Python, Databricks

3. Reporting

- a. Turning the results of validation/analysis into displayable insights (i.e., Tableau, PowerBI, Excel Pivot)
- b. Providing management with presentations, an overview of our pipelines, teams involved, upcoming milestones, weaknesses/strengths of the current process, and areas for improvement.
- c. Reports tracking other technical aspects of the pipeline, such as preprocessing transformations (ensuring they're working), market transparency reports of competitors, and firms where we're lacking transparency (i.e., access to identifiable data)

4. Documentation

- a. A well-maintained catalog of all technical, crucial aspects of ETF Data here at Capital Group.
- b. Used for training.
- c. Referenced for other technical and business teams.
- d. Crucial for change management events