

Technology Forum - Open Data and Analytics

Unleash Data with Alteryx

Presenter: Jan Chan



FASTER ANSWERS. MORE INSIGHTS. BETTER OUTCOMES.

Velocity Business Solutions Limited © 2010-2020, Confidential. www.vebuso.com



LATE NIGHTS. LONG HOURS.

Due to many analytic inefficiencies, employees are stuck in spreadsheet hell trying to analyze their data — instead of using their time wisely to affect changes based on what the data reveals.

\$6
BILLION

hours per year are spent working in spreadsheets.

26
HOURS

per week are wasted working in spreadsheets.

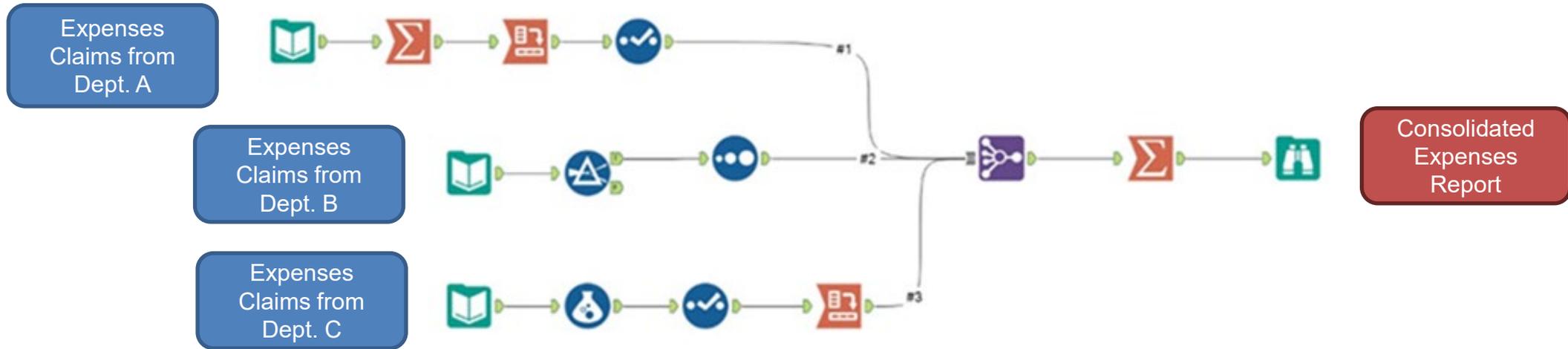
8
HOURS

per week wasted repeating the same data tasks.

Source: "The State of Self-Service Data Preparation and Analysis Using Spreadsheets," IDC



Alteryx Workflow Approach is Superior



Consistent

Preparation, analytics and reporting in the same workflow



Transparent & Self-Documenting

See data flow through the process, and the actions performed at every step



Flexible

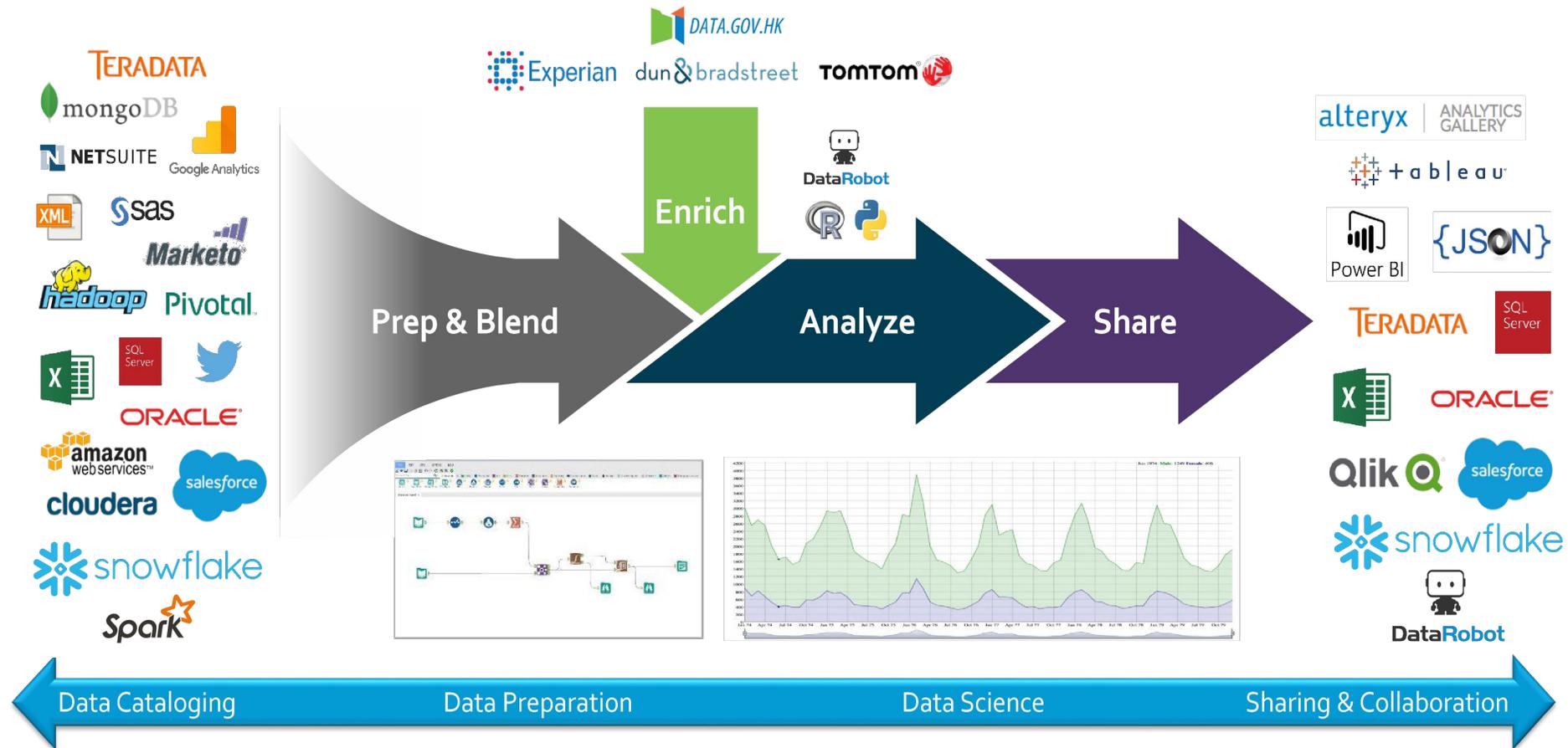
Supports more than just structured data that fits nicely into rows and columns



Repeatable

Re-run (manually or automatically) whenever you need updated results

Alteryx: Analytic Process Automation (APA) Platform For Citizen Data Scientist



[Alteryx Supported Data Sources](#)



Workflow-based Approach is Superior



Consistent

Preparation, analytics and reporting in the same workflow



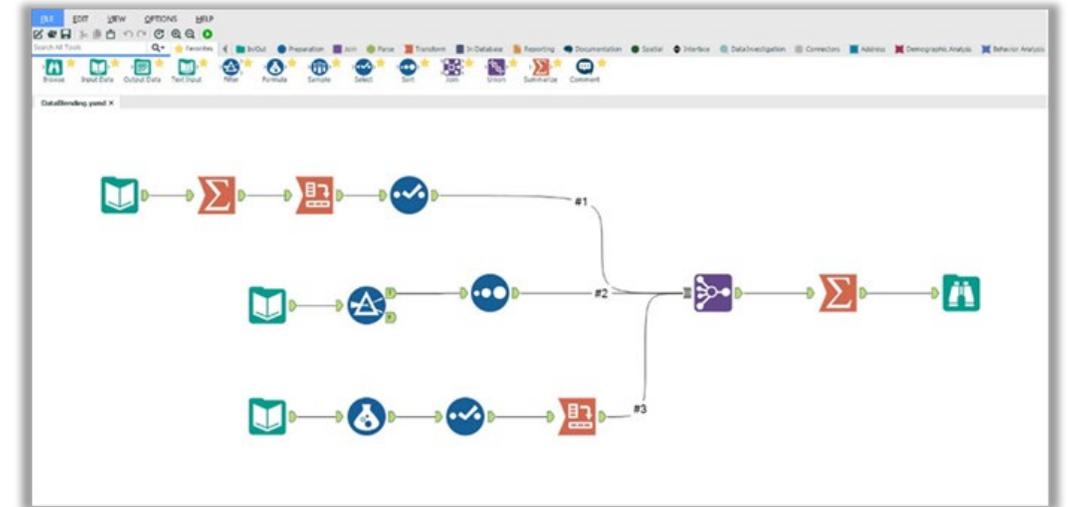
Transparent & Self-Documenting

See data flow through the process, and the actions performed at every step



Flexible

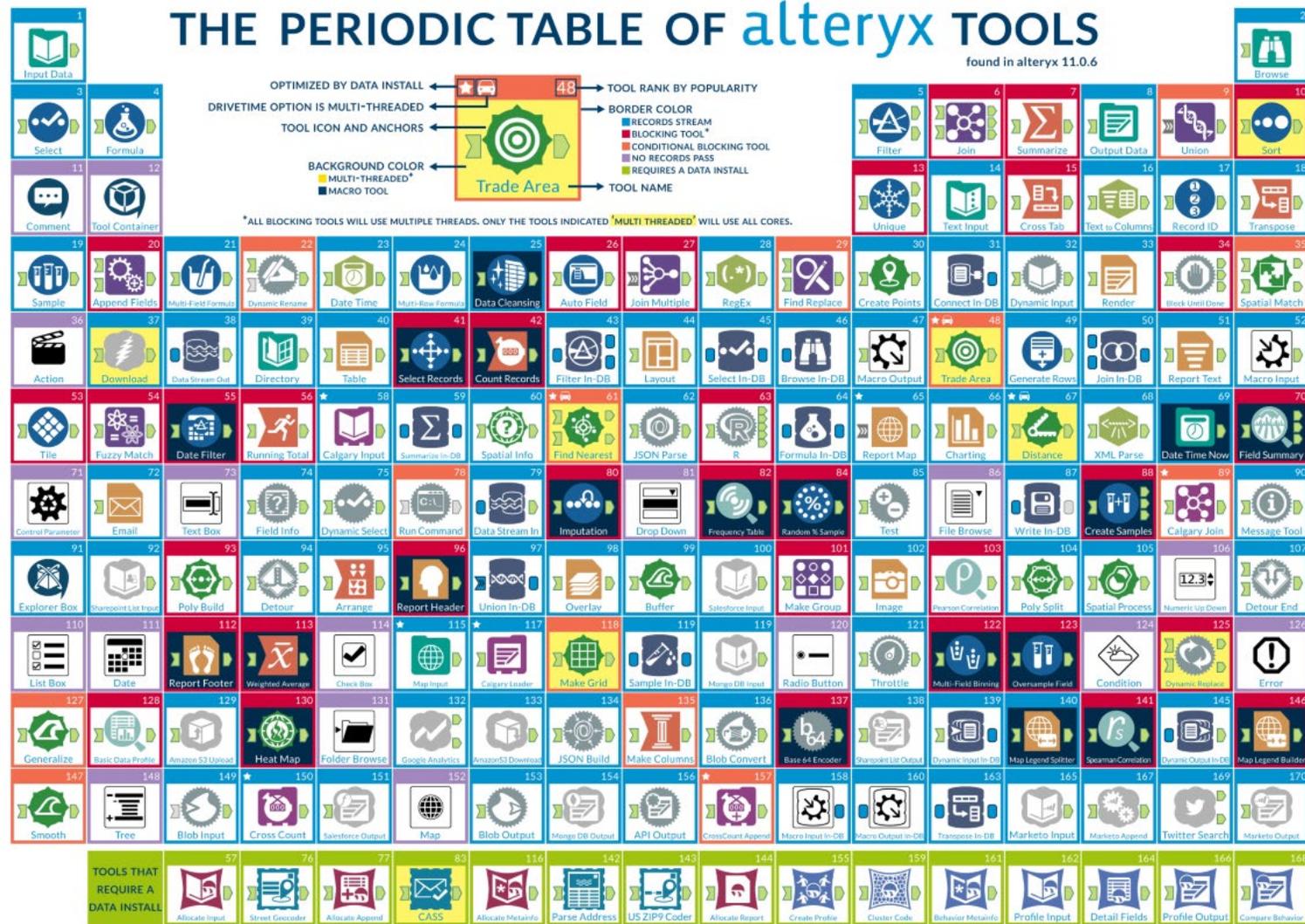
Supports more than just structured data that fits nicely into rows and columns



Repeatable

Re-run (manually or automatically) whenever you need updated results

Over 250+ Tools



Collections of Advanced Analytics Tools



THE PREDICTIVE TOOLS OF **alteryx**

found in the alteryx 11.0.6 predictive install

Descriptive Tools

WHAT'S HAPPENED?

Association Analysis	Plot of Means	Append Cluster	MB Inspect
Contingency Table	Scatterplot	Find Nearest Neighbors	MB Rules
Distribution Analysis	Violin Plot	K-Centerable Analysis	Principal Components
Heat Plot		K-Centerable Diagnostics	
Histogram		MB Affinity	

Predictive Tools

WHAT'S GOING TO HAPPEN?

Boosted Model	Lift Chart	Network Analysis	Support Vector Machine	ARIMA	TS Forecast
Count Regression	Linear Regression	Neural Network	Test of Means	ETS	TS Plot
Decision Tree	Logistic Regression	Score		TS Compare	
Forest Model	Naive Bayes Classifier	Spline Model		TS Covariate Forecast	
Gamma Regression	Nested Test	Stepwise		TS Filler	

Prescriptive Tools

WHAT'S THE BEST THING TO DO?

AB Analysis	AB Controls	Optimization	Simulation Scenarios
AB Treatments	AB Trend	Simulation Scoring	Simulation Summary

THE PREDICTIVE TOOLS ARE MACROS.
TOOLS DISPLAYING A BLUE BACKGROUND
HAVE AN HTML FRONT END.

PREDICTIVE MODEL CHEAT SHEET

Predictive Modeling coming from statistics

Fast and intuitive but tend to make assumptions about the underlying data, tend to be oriented around hypothesis testing

Machine Learning coming from computer science

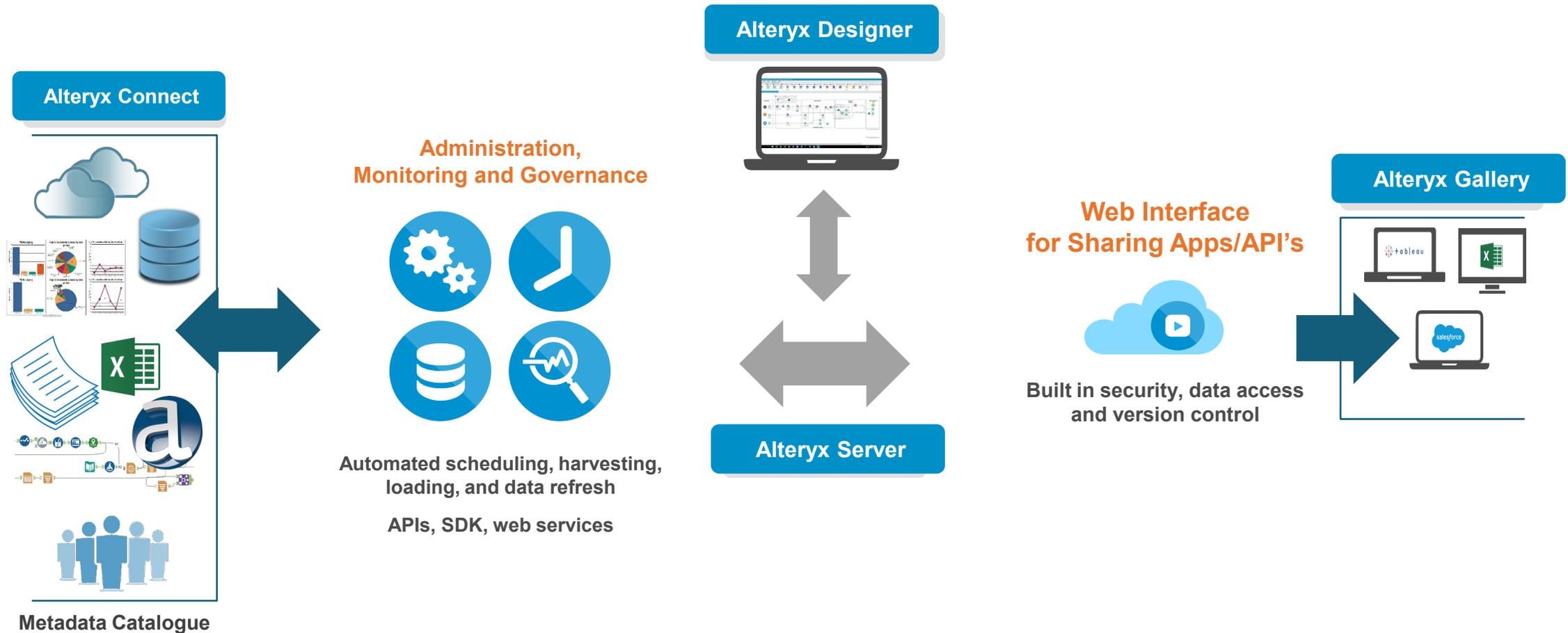
Algorithms can be more accurate but tend to take longer to fit and are harder to interpret

Regression			Regression & Classification					
Classification								

Regression = Predict continuous numbers | Classification = Predict categories



The Alteryx Platform

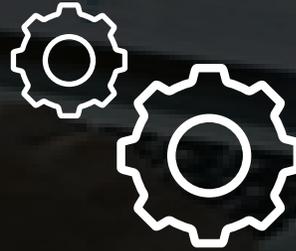




Use Cases



ETL Use Case: Regional Finance Dashboard



Global Insurance Company

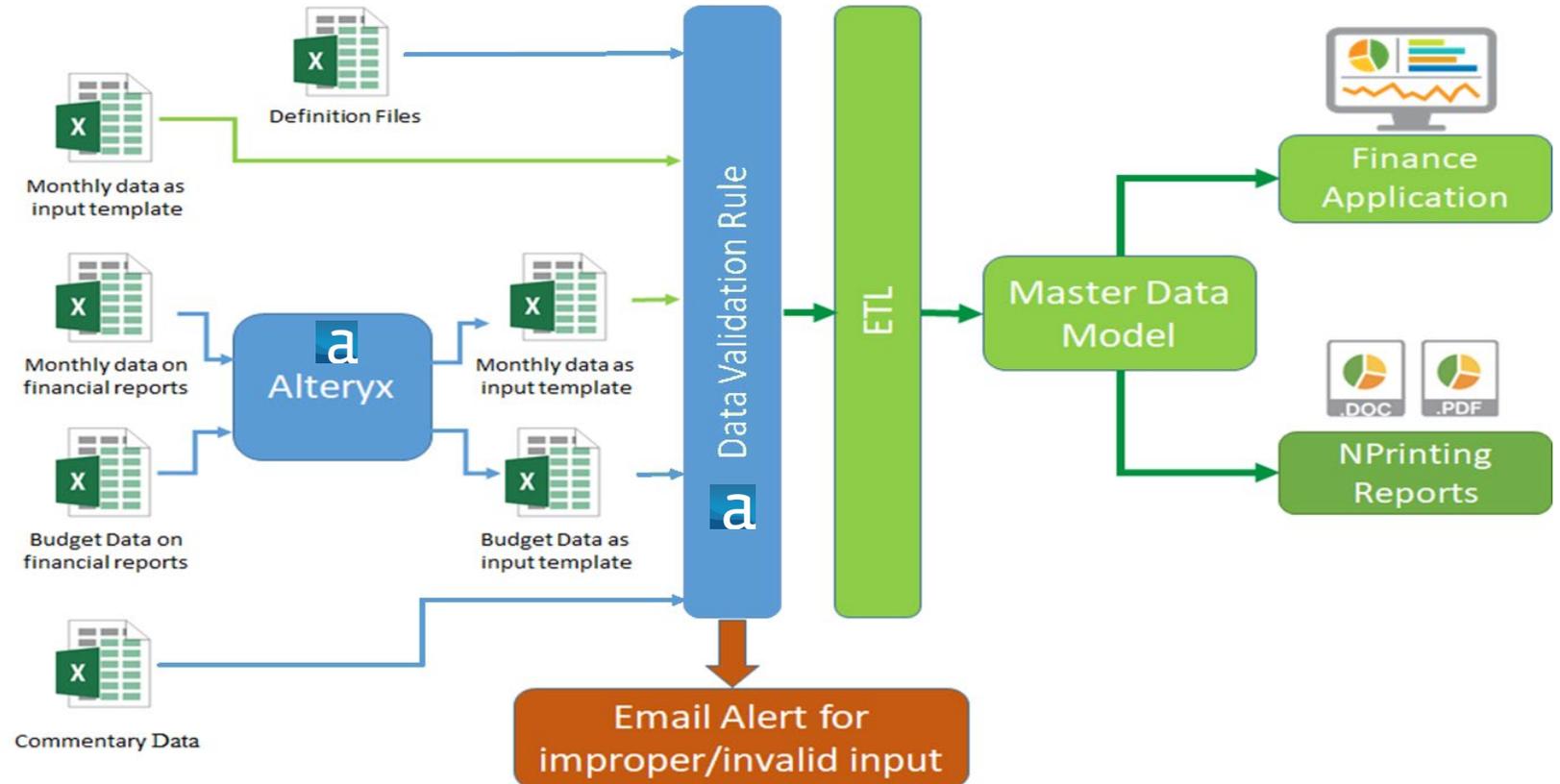
Department: Finance

Business Problem:

Require at least 2 weeks to prepare the reports for management from extraction, validation, and visualization

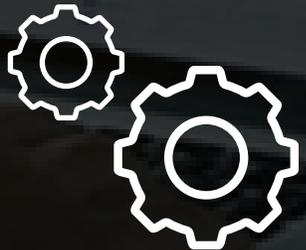
Benefits:

- ✓ Time saved: days vs weeks
- ✓ Reduced risk: manual eyeball checking vs automated validation rules and processes
- ✓ Improved sales opportunities





ETL Use Case: Autopay Reconciliation



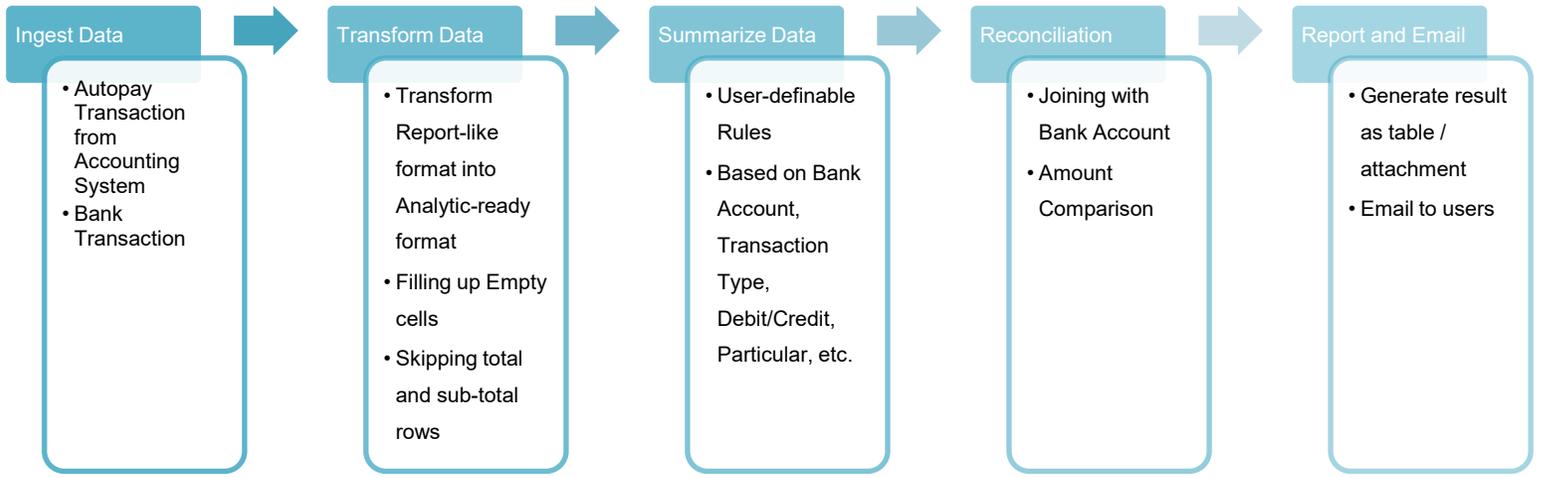
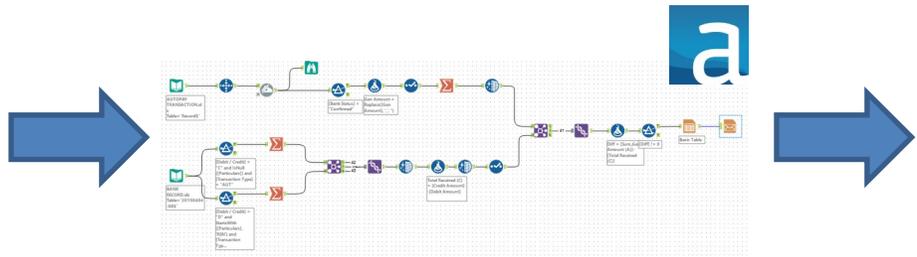
Property
Management
Service
Company

Department: Finance

Business Problem: Repetitive data transformation applied to individual Bank Accounts with varies business rules based on account type.

Benefits:

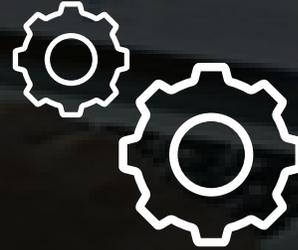
- ✓ Reduce Error by Automating the process for reconciliation for large amount of records across multiple bank accounts
- ✓ Flexible to adopt new business rules
- ✓ Alert user with exception through email
- ✓ Provide transparency to business users

Bank Account	Sum_Gen Amount (A)	Total Received (C)	Diff
693640	667550	26090	
588200	572410	15790	
131404	122724	8680	
766707	729984	36723	
368218	362400	5818	
704400	694154	10246	
486589	446524	20065	
2184781	2119004	64777	
469712	457233	12479	
150234	145564	4670	
767901	754924	12977	
121082	120399	683	
145360	134320	11040	



Sales ETL Use Case: 360° Customer Dashboard



3D Printer
Manufacturer

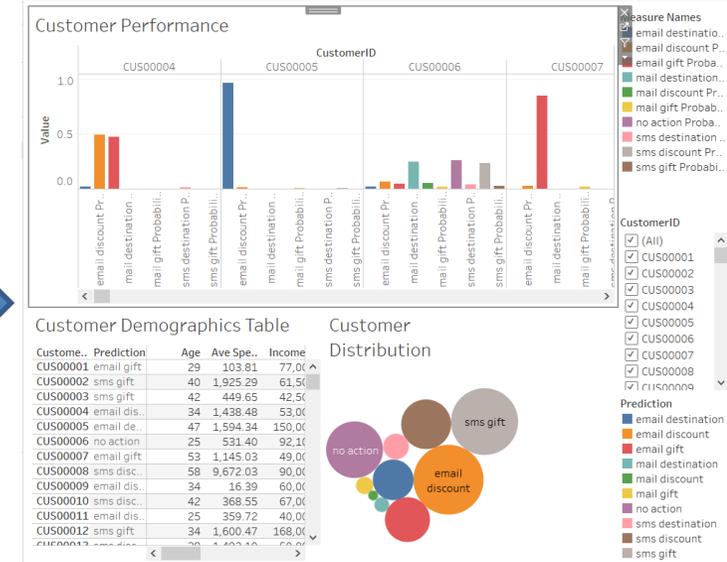
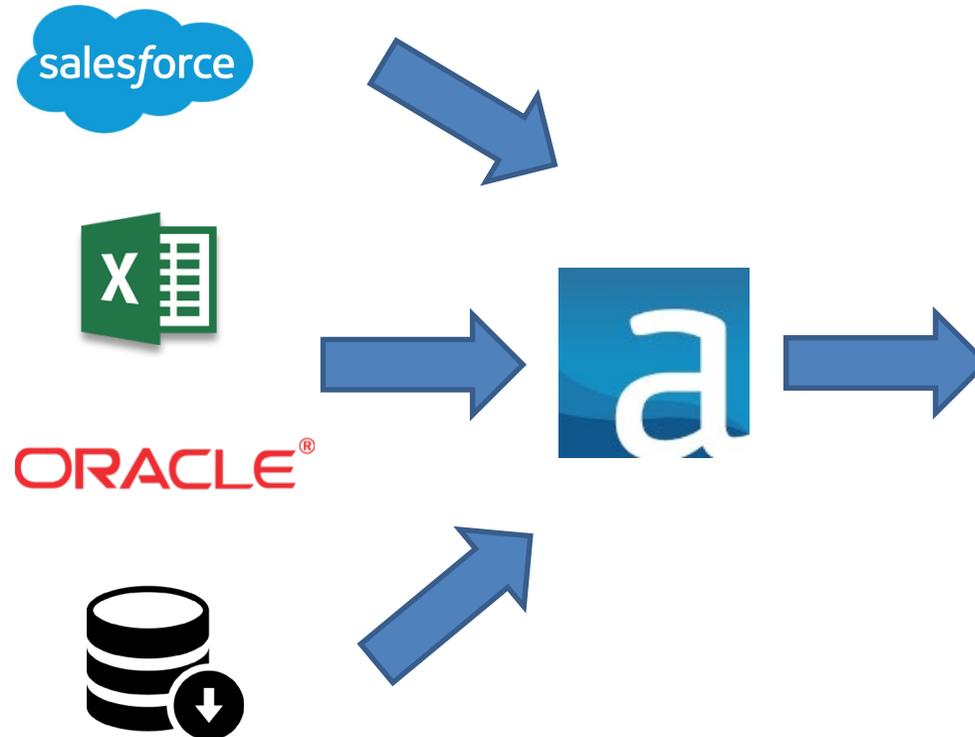
Department: Sales

Business Problem:

- Increasing number of data sources: Salesforce, Excel, Oracle etc.
- Data sizes are huge for Excel to handle
- Varying data formats per region
- Even with visualization tool, the marketing dashboard still needs 5 hours to generate

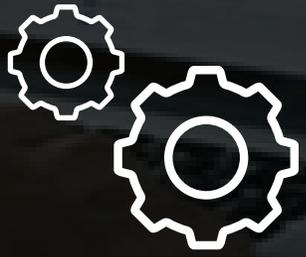
Benefits:

- ✓ Efficiency Gains and Big Time Savings: 1 Day vs 30 mins
- ✓ More Effective Communications: Better documentation and collaboration between departments





ETL Use Case: Form Analysis (Semi-structured to Structured)



Department: Human Resources

Business Problem: Time consuming to extract claim form detail for analysis

- Benefits:**
- ✓ Streamline the whole process in one single run: Read and validate all PDF files in a folder and prepare data for dashboard
 - ✓ Efficiency improved 80%



Insurance Claim Form PDF

Step 1



Text format

Step 2



Staff No.	Name of Insured	Hospital	Admission Date	Discharge Date
AC013	Justin Lee	SAC Hospital	2017-09-03	2017-09-10
IT085	Natalie Cheung	St. Mary's Hospital	2017-09-15	2017-09-21
..

Step 3



Reply | Reply All | Forward

The 4/8/2017 12:19 PM

jasperich@gmail.com

To: Jasper Lee

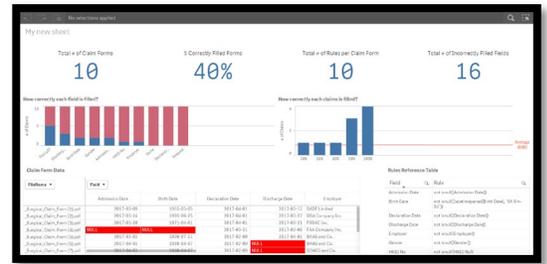
2 KB

New! In table.csv

Correct Incorrect

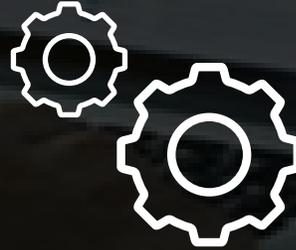
001_Check_PolicyID	5
002_Check_Name	10
003_Check_Birth_Date	8
004_Check_Gender	8
005_Check_Declaration_Date	10
006_Check_HKID_No	9
007_Check_Employer	10
008_Check_Hospital	10
009_Check_Admission_Date	8
010_Check_Discharge_Date	7

Generate Emails for validation result and Data for visualization





Web Scraping Case: Generate Stock Insider News for Clients

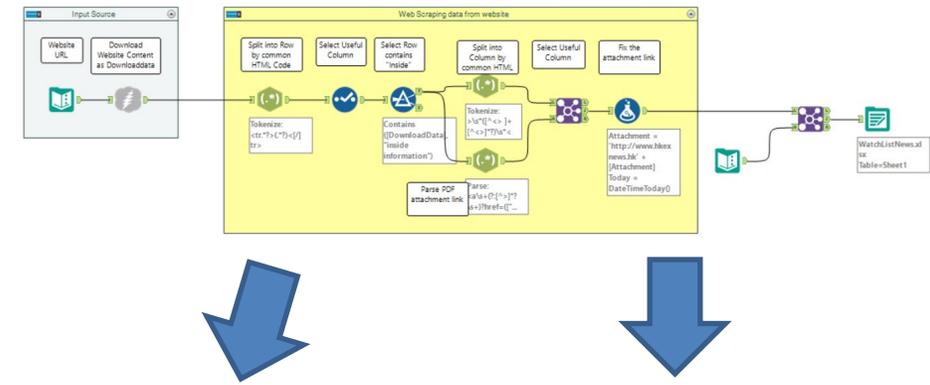
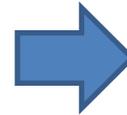
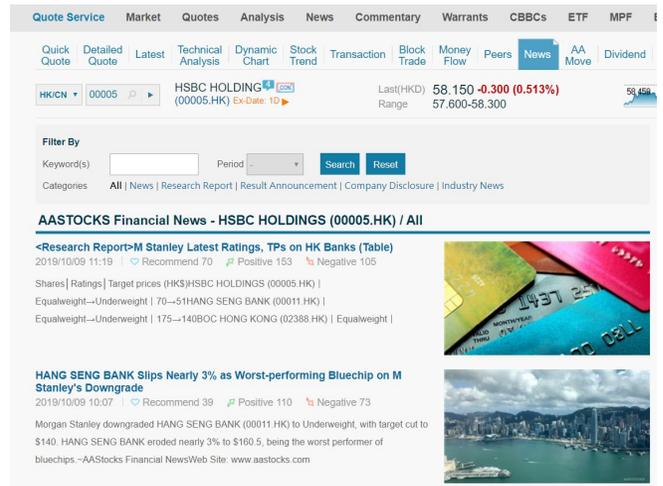


Investment Bank

Department: Brokerage

Business Problem: Need to spend at least an hour for manually grabbing day-end prices and stock news from website daily

Benefits: Only need to press Run daily now and automated reports can be generated within 5 minutes

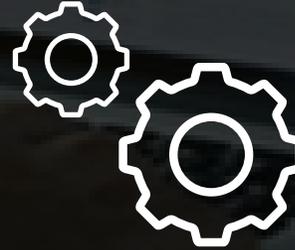


Record	Stock	Titles	DateTime	Contents	Dates	52Wk_High	52Wk_Low	Ask	Bid
1	00005	<Research Report> M Stanley Latest Ratings, T...	2019/10/09 11:19	Shares/Ratings/Target prices (HK\$)HSBC HOLDI...	2019-10-09	70.50	55.30	58.10	58.05
2	00005	HANG SENG BANK Slips Nearly 3% as Worst-p...	2019/10/09 10:07	Morgan Stanley downgraded HANG SENG BA...	2019-10-09	70.50	55.30	58.10	58.05
3	00005	<Research Report> M Stanley Cuts TPs on HK...	2019/10/09 10:05	Morgan Stanley did not expect large banks to...	2019-10-09	70.50	55.30	58.10	58.05
4	00005	'M Stanley Cuts TPs on HK Banks; Downgrades...	2019/10/09 10:02	--AAStocks Financial NewsWeb Site: www.aast...	2019-10-09	70.50	55.30	58.10	58.05
5	00700	TENCENT Reportedly Splashes US\$500M in Hill...	2019/10/09 12:53	TENCENT (00700.HK) is deliberating to invest...	2019-10-09	400.40	251.40	319.40	319.20
6	00700	Tencent 'Call of Duty: Mobile' Picks Up Over 10...	2019/10/09 12:11	TENCENT (00700.HK)'s Call of Duty: Mobile ge...	2019-10-09	400.40	251.40	319.40	319.20
7	00700	Sensor Tower: TENCENT's PUBG Mobile Tops G...	2019/10/09 10:36	TENCENT (00700.HK)'s PUBG Mobile topped L...	2019-10-09	400.40	251.40	319.40	319.20
8	00999	[Null]	[Null]	[Null]	[Null]	7.21	5.55	6.01	5.99

(00005) HSBC HOLDINGS <Research Report> M Stanley Latest Ratings, TPs on HK Banks (Table) Shares/Ratings/Target prices (HK\$)HSBC HOLDINGS (00005.HK)Equalweight--Underweight 70--51HANG SENG BANK (00011.HK) Equalweight--Underweight 175--140BOC HONG KONG (02388.HK) Equalweight&# Ask:58.10 Bid:58.05 Day High:58.30 Day Low:57.60 Market Cap.:1,203.36B Volume:22.68M Turnover:1.32B P/E:11.71
(00005) HSBC HOLDINGS <Research Report> M Stanley Cuts TPs on HK Banks; Downgrades HSBC, HSB to Underweight Morgan Stanley downgraded HANG SENG BANK (00011.HK) to Underweight, with target cut to \$140. HANG SENG BANK eroded nearly 3% to \$160.5, being the worst performer of bluechips.--AAStocks Financial NewsWeb Site: www.aastocks.com Ask:58.10 Bid:58.05 Day High:58.30 Day Low:57.60 Market Cap.:1,203.36B Volume:22.68M Turnover:1.32B P/E:11.71
(00005) HSBC HOLDINGS <Research Report> M Stanley Cuts TPs on HK Banks; Downgrades HSBC, HSB to Underweight Morgan Stanley did not expect large banks to maintain their average ROCET1 level at over 20% under a low interest rate environment, economic growth slowdown and market competition. The outlook for Hong Kong's banking industry was downgraded to Cautio Ask:58.10 Bid:58.05 Day High:58.30 Day Low:57.60 Market Cap.:1,203.36B Volume:22.68M Turnover:1.32B P/E:11.71



Advanced Analytics Case: Forecast on No. of Passengers

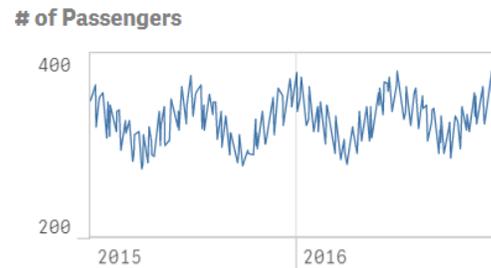


Airline Holding Company

Department: Strategic Planning

Business Problem: Improve the allocation of resources

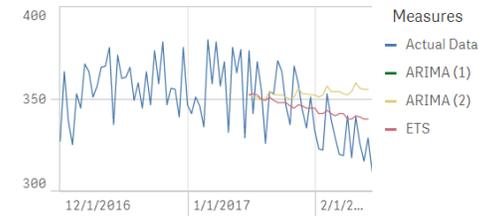
Benefits: Business users can now do forecasting and prediction code-free without lots of help from other teams



Gather historical data



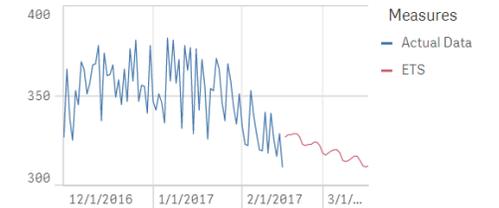
of Passengers (Actual vs Prediction)



Run multiple predictive models and compare the results with actual data



of Passengers (Prediction)



Choose the best model and make future predictions

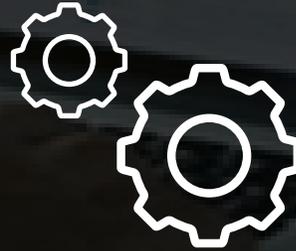


Visualization in Qlik Sense





Marketing Use Case: Extract online stores and social media data to Data Warehouse



Property
Development

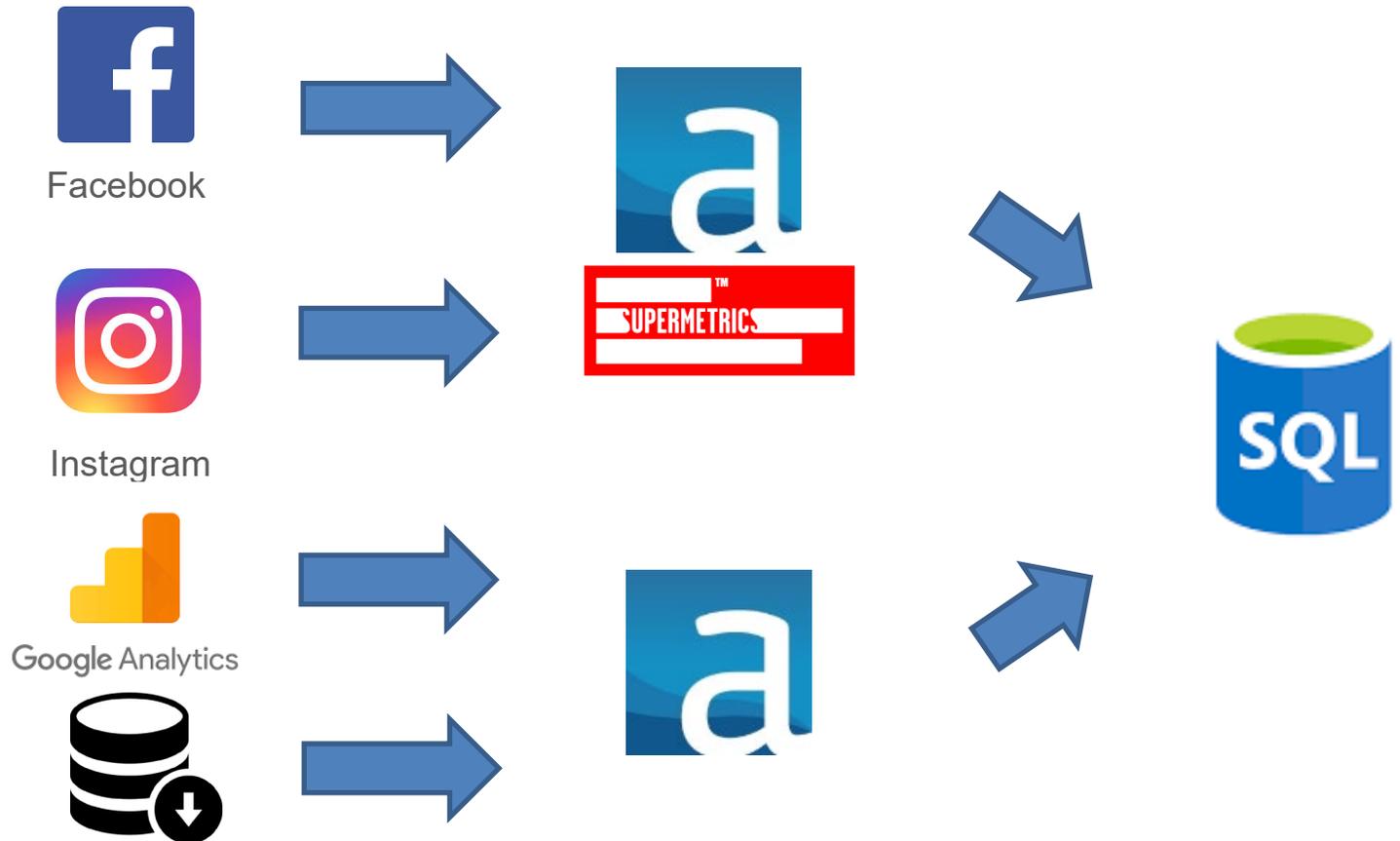
Department: Marketing & Digital

Business Problem:

- Social media data are only captured in Google Sheets through Supermetrics
- Difficult to blend with their other data

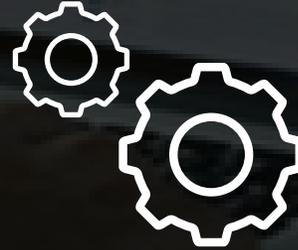
Benefits:

- ✓ All social media data are consolidated in SQL through Supermetrics API
- ✓ Developing a marketing dashboard with internal and social media data





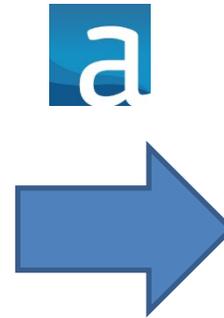
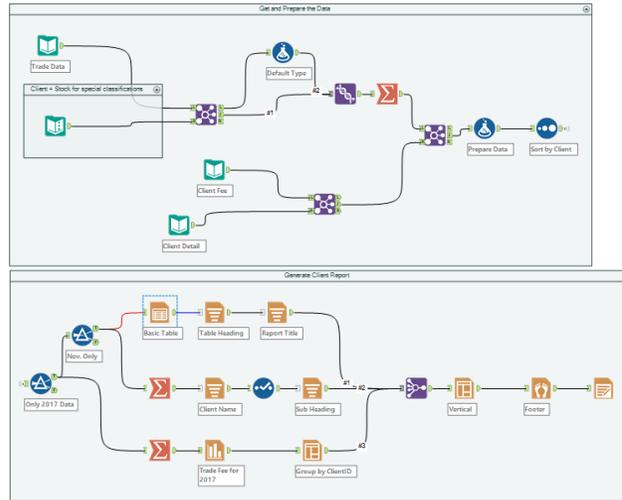
Reporting Case: Generate Monthly Client Trade Report



Department: Customer Support

Business Problem: Streamline the manual process of sending monthly report to client

- Benefits:**
- ✓ Time saved over 90%
 - ✓ Emails can be sent automatically to client with encryption



For each client, a PDF with encryption is sent with the following sample layout

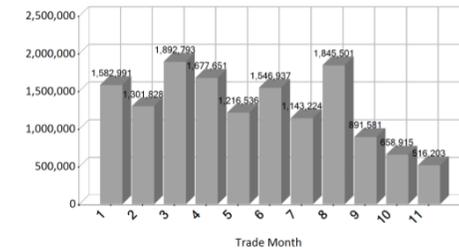
Monthly Client Trade Report Barclays Capital Inc.

Total Trade Fees for Nov, 2017 is \$516,203

Monthly Trading Volume and Fees

Date	No. of Trade	Trade Volume	Trade Value	Trade Fee
2017-11-01	4	207,523	9,260,777	37,043
2017-11-02	1	24,721	847,270	3,368
2017-11-03	1	77,225	4,955,614	12,414
2017-11-03	2	159,524	11,396,707	45,595
2017-11-04	5	400,036	28,713,041	114,852
2017-11-05	7	494,070	33,420,852	133,663
2017-11-09	4	313,875	14,297,407	57,190
2017-11-10	1	86,970	10,635,474	42,542
2017-11-11	4	230,498	9,960,326	39,841
2017-11-12	2	159,706	7,413,336	29,653

Current Year-to-Date Trading Fees



Alteryx Hourly Series - Reporting in Alteryx for Begin

Name	D.
BAR1_2017_11.pdf	11
CIT1_2017_11.pdf	11
CRE1_2017_11.pdf	11
DEU1_2017_11.pdf	11
GOL1_2017_11.pdf	11
JPM1_2017_11.pdf	11
MER1_2017_11.pdf	11
MOR1_2017_11.pdf	11
UBS1_2017_11.pdf	11

Thank You

<https://www.vebuso.com/contact-us/>

