

Analytics mindset

ETL

Case 1 – Identifying data problems – Alteryx

To successfully extract, transform and load (ETL) data, you need to be able to identify potential data problems and challenges. The image below contains balance sheet data for several different companies. The image shows a screenshot of the comma-separated file provided to you for this case.

File	Edit	Format	View	Help
Ticker,Name,Year,Cash,Accounts Receivables,Inventories,Total Current Assets,Total Depreciation				
AAL	American Airlines Group Inc.	,2016	3220000000,"1,600,000,000"	1094000000,10300,(14200000000)
AAL	American Airlines Group Inc.	,2017	2950000000,"1,800,000,000"	1359000000,9100,(15600000000)
AAL	American Airlines Group Inc.	,2018	2750000000,"1,700,000,000"	1522000000,8600,(17400000000)
AAPL	Apple, Inc.	,2016	20500 M,"15,800,000,000"	2100000000,106900,(34200000000)
AAPL	Apple, Inc.	,2017	20300 M,"17,900,000,000"	4900000000,128600,(41300000000)
AAPL	Apple, Inc.	,2018	25900 M,"23,200,000,000"	4000000000,131300,(49100000000)
ACY	Aerocentury Corp.	,2016	2200000,4000000,0,6,-23000000	
ACY	Aerocentury Corp.	,2017	8700000,3800000,-,12,-14600000	
ACY	Aerocentury Corp.	,2018	1500000,4000000,0,5,-14700000	
AEHR	Aehr Test Systems	,2016	900 K,"500,000"	7000000,9,-5700000
AEHR	Aehr Test Systems	,2017	17800 K,"4,000,000"	6600000,29,-6000000
AEHR	Aehr Test Systems	,2018	16800 K,"2,900,000"	9000000,29,-6400000
	Armada Hoffler Properties Inc.	,2016	21900000,"15,100,000",,0,-139600000	
	Armada Hoffler Properties Inc.	,2017	19900000,"15,700,000",,0,-164500000	
	Armada Hoffler Properties Inc.	,2018	21300000,"19,100,000",- ,0,-188800000	
AMZN	Amazon.com Inc.	,2016	19330000000,8340000000,11500000000,45780,(13330000000)	
A M Z N	Amazon.com Inc.	,2017	31750000000,16680000000,17170000000,162650,(33970000000)	
AM ZN	Amazon.com Inc.	,2018	20520000000,13160000000,16050000000,60200,(23790000000)	
APVO	Aptevo Therapeutics Inc.	,2016	9700 K,"4300000",6600000,71,(6600000)	
APVO	Aptevo Therapeutics Inc.	,2017	7100 K,"2100000",1000000,95,(7500000)	
APVO	Aptevo Therapeutics Inc.	,2018	30600 K,"5200000",1800000,49,(8700000)	
APVO	Aptevo Therapeutics Inc.	,2018	30600 K,"5200000",1800000,49,(8700000)	

Required

- ▶ Complete the ETL overview case, which covers fundamental considerations in the ETL process.
- ▶ Review the data extract shown above. Identify any issues that might be problematic and would need to be addressed when performing ETL procedures.
 - * Note: You do not need to verify the validity of the amounts (e.g., American Airlines Group Inc. did in fact report cash of 2750000000 for 2018).
- ▶ Transform the provided data using Alteryx to address the issues you identified. Produce an Alteryx Database (.yxdb) output file at the end of the workflow.
- ▶ Submit your Alteryx workflow and output database file as a packaged workflow (.yxzp file type [Options > Export Workflow >]) named as follows: **LastnameFirstname Alteryx Assessment.yxzp**.