Released 2021-Apr-16 Revised 2021-Apr-20 (Question 4, Submission)

TractorTEK has recently reconstructed their Sales Decision Support System, in record time by any measure, with the aid of capable TEKsystems Data Analytics consultants. Determined to evolve the current system into a full-scale Enterprise Analytics Platform, they've become repeat customers and are seeking a longer-term strategy going forward from the same team.

Top of mind is how to level up their data processing capabilities in a way that scales. This means finding a tool set that integrates with a wide range of disparate source systems and reporting systems. They intend to connect to SQL Databases, Excel files, and various text file repositories; then process the data, and expose the data to reporting tools such as Tableau and Power BI. Most department analyst teams have non-coders used to Excel and some SQL. A couple teams have some data scientists that use the usual toolsets (Python, R).

You're tasked with preparing a very basic proof of concept for a tool that's right in TEK's wheelhouse: **Alteryx**. Using the dataset previously used for the Capstone, and the Subject Matter Expertise bestowed upon you by guest lecturer Michael Perillo and Udemy, please do the following exercises:

1. **Demo** how easy it is to process data. Take each of the raw Capstone datasets *directly from* Excel, and process them into a CSV format that's suitable for ingestion to a database or by any Data Visualization tool. Do so by recreating this table structure in Alteryx. **Show your work** with clearly labeled workflow steps and comments.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	*(includes prod codes & esp codes)*																
2	ItemCode	Sales Year	W#	Week Number	EmpID	Qty	Price	Total$	Sales Period	Sales Quarter	Week Start Date	Week End Date	Employee Paygrade	Region	Manufacturer	Firstname	Lastname
3	PROD_001	2019	W0	1	EMP244	35	12449	435715	1	2019Q1	12/30/2018	1/5/2019	C12	NW	John Deere	Gina	Evans
4																	

2. Building off of that demo, explain in plain English the ways in which Alteryx may streamline a company's ability to do large scale data processing quickly. Provide no more than 2-3 paragraphs tops. Factors should include but are not limited to:
    a. The comparable effort it would take to build a SQL database that does the same thing.
    b. Accessibility by coders and non-coders, in contrast to Excel and Python.
    c. Auditability & traceability. How can Alteryx facilitate trust internally and externally (government auditors) that data processing is reliable and transparent?

3. Build upon the first demo: relocate ESP codes so that ItemCode column now only has Prod Codes, and there are new columns ESPCode, ESP Qty, ESP Price, ESP Totals as follows:
   a. 

R	S	T	U
ESPCode	ESP Qty	ESP Price	ESP_Total$
ESP_001	9	843	7587

Problem 4 is cancelled because it's not reasonably easy to achieve as written. You're encouraged to look into the Reporting section of the Udemy course to understand how it works, but are not required to submit any work for HW4.

4. Develop a basic data visualization in Alteryx with a header, and two static tables (no need for dynamic buttons):
   a. Sales Attachment Rates by Region
   b. Sales Attachment Rates by Employee broken into two categories: "Top 2 Performers" and "Bottom 3 Performers". This will require that you create some additional fields from Step 3.

Submission: Your submission should include an Alteryx file and a Word file each with naming scheme "hw4_lastname", saved to a Github repo. *The Word file must include your response to Question 2, but it may also include Question 1's "show your work" if you feel that naming the steps and adding comments in Alteryx don't adequately explain your work.*