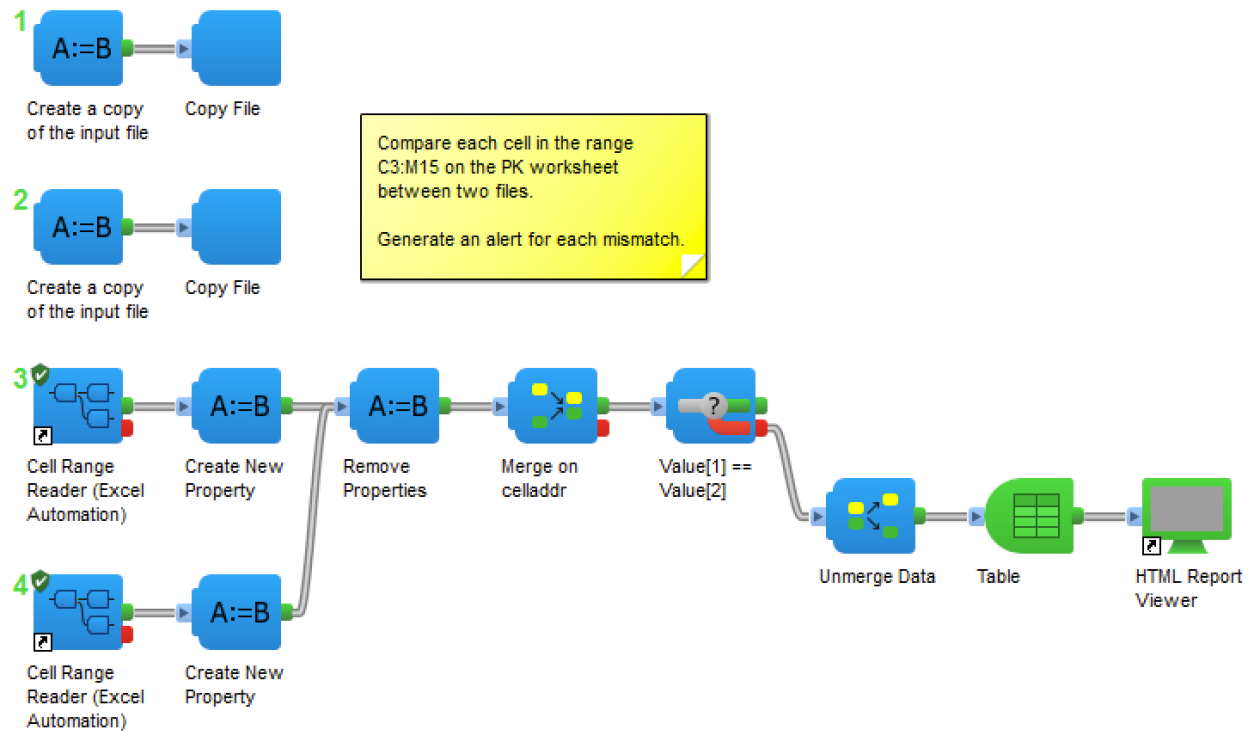


09 Reconcile Data Between Two Files

Purpose

Automatically match values for each cell in the same cell range in two different Excel files, and identify discrepancies. Generate an alert for each mismatch.



Workflow

In this example we compare each cell in the range C3:M15 on the **PK** worksheet between two files: data/Animal Safety Study.xlsx and data/Animal Safety Study v2 XLSformat.xlsx.

Animal ID	Treatment	1	2	3	4	5	6	7	8	9	10	11
1	83.07826346	154.8818471	159.3096813	151.22749	171.829946	117.3759036	153.674108	155.4314726	156.5453634	111.9631173	119.5541901	
2	31.58151871	39.39974838	21.69388553	0.068195632	27.50303359	34.03315317	42.15472871	26.72434541	22.46703628	12.21921277		
3	36.08318228	35.30618701	35.38526436	23.13515272	22.98309965	34.16594163	31.77837933	40.14941419	29.61079578	26.49642763	19.16799174	
4	31.53094983	34.83854911	31.12939676	18.74200499	22.02759859	33.43518321	28.33263003	31.59917652	31.05877289	20.6566854	15.48829358	
5	26.91945066	32.14973856	18.35007736	15.34265574	18.96520304	23.28647037	30.9392931	26.44686947	24.87112885	17.62031577	12.23831351	
6	38.45292607	43.40300704	21.64522857	20.44680744	23.04272192	28.85859868	36.2601067	24.48752207	41.2855801	17.34188452	12.36566363	
7	29.51585633	40.48714908	19.86088718	17.373167	15.72850459	27.23409045	37.88719814	23.93720918	34.28028114	17.28025078	10.36531244	
8	26.28635011	30.83712568	19.96997547	17.36745186	16.01050228	20.89576329	34.58604182	17.4882432	24.30792209	13.78954466	11.5047396	
9	27.52415366	29.65451477	14.45247803	15.52979034	13.86400587	6.957439222	25.86001334	15.60063529	20.64070655	16.82507316	12.33825044	
10	33.5646725	22.488376	16.05831948	15.28607504	16.07071661	20.52726809	23.77062511	13.7468168	19.88918603	15.85552971	10.76574147	
11	21.06712498	21.36536988	11.37463147	14.72898121	11.964217	16.96916799	18.74439975	10.0315567	16.69278342	15.061214	35.76956179	
12	26.55231788	24.91403988	8.394792878	24.38134743	11.27013151	13.41128195	21.89548103	10.42691593	22.29640387	22.58262542	3.626903342	
13	32.52877807	19.19385454	11.48773688	32.56999167	8.335339616	13.63622782	24.02813574	8.030024376	16.16827553	30.35935017	7.491231584	
14	AUC	378.1	440.12	302.1	275.34	279.2	328.84	419.23	320.12	378.11	269.08	279.56

Animal ID	Treatment	1	2	3	4	5	6	7	8	9	10	11
1	83.07826346	154.8818471	159.3096813	151.22749	171.829946	117.3759036	153.674108	155.4314726	156.5453634	111.9631173	119.5541901	
2	31.58151871	39.39974838	21.69388553	0.068195632	27.50303359	34.03315317	42.15472871	26.72434541	22.46703628	12.21921277		
3	36.08318228	35.30618701	35.38526436	23.13515272	22.98309965	34.16594163	31.77837933	40.14941419	29.61079578	26.49642763	19.16799174	
4	31.53094983	34.83854911	31.12939676	18.74200499	22.02759859	33.43518321	28.33263003	31.59917652	31.05877289	20.6566854	15.48829358	
5	26.91945066	32.14973856	18.35007736	15.34265574	18.96520304	23.28647037	30.9392931	26.44686947	24.87112885	17.62031577	12.23831351	
6	38.45292607	43.40300704	21.64522857	20.44680744	23.04272192	28.85859868	36.2601067	24.48752207	41.2855801	17.34188452	12.36566363	
7	29.51585633	40.48714908	19.86088718	17.373167	15.72850459	27.23409045	37.88719814	23.93720918	34.28028114	17.28025078	10.36531244	
8	26.28635011	30.83712568	19.96997547	17.36745186	16.01050228	20.89576329	34.58604182	17.4882432	24.30792209	13.78954466	11.5047396	
9	27.52415366	29.65451477	14.45247803	15.52979034	13.86400587	6.957439222	25.86001334	15.60063529	20.64070655	16.82507316	12.33825044	
10	33.5646725	22.488376	16.05831948	15.28607504	16.07071661	20.52726809	23.77062511	13.7468168	19.88918603	15.85552971	10.76574147	
11	21.06712498	21.36536988	11.37463147	14.72898121	11.964217	16.96916799	18.74439975	10.0315567	16.69278342	15.061214	35.76956179	
12	26.55231788	24.91403988	8.394792878	24.38134743	11.27013151	13.41128195	21.89548103	10.42691593	22.29640387	22.58262542	3.626903342	
13	32.52877807	19.19385454	11.48773688	32.56999167	8.335339616	13.63622782	24.02813574	8.030024376	16.16827553	30.35935017	7.491231584	
14	AUC	378.1	440.12	302.1	275.34	279.2	328.84	419.23	320.12	378.11	269.08	279.56

Before reading any data we make a copy of both input files in the job directory. The original input files are shared by all users of the exercise: creating a copy helps ensure the original is not accidentally modified if you later decide to use a writer component on the file.

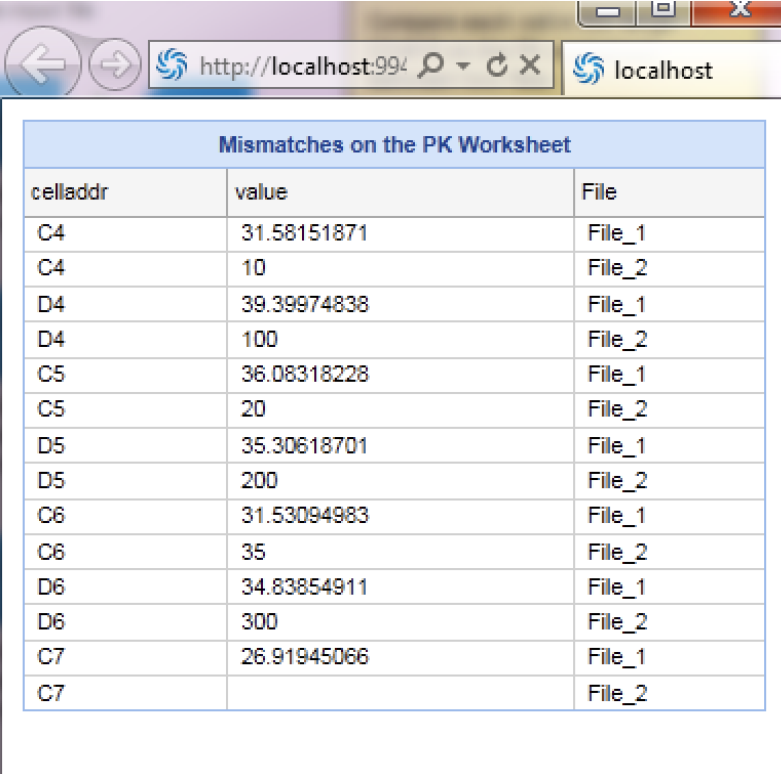
Using the **Cell Range Reader** component, we read from our copy of each file and extract the values from the C3:M15 range on the **PK** worksheet. We send them into Pipeline Pilot as data records with the property "value".

We merge cell data on cell address (e.g. C8, E10, etc.) so that each cell pair becomes an array of values - as we are reading from two files.

Using a PilotScript filter component we compare the values in each pair and if they differ or one is missing, generate an alert in a web browser window.

Results

Here is the result of the protocol.



The screenshot shows a web browser window with the address bar displaying 'http://localhost:994' and the page title 'localhost'. The main content area contains a table with the title 'Mismatches on the PK Worksheet'. The table has three columns: 'celladdr', 'value', and 'File'. It lists 16 rows of data, each representing a mismatch between two files (File_1 and File_2) for a specific cell address and value.

Mismatches on the PK Worksheet		
celladdr	value	File
C4	31.58151871	File_1
C4	10	File_2
D4	39.39974838	File_1
D4	100	File_2
C5	36.08318228	File_1
C5	20	File_2
D5	35.30618701	File_1
D5	200	File_2
C6	31.53094983	File_1
C6	35	File_2
D6	34.83854911	File_1
D6	300	File_2
C7	26.91945066	File_1
C7		File_2