Data Wrangling Problem on a Retail Data

Misbah Uddin, Jun Xie

The following dataset provides information on weekly sales from stores of a retail chain

sales

	Store	Dept	Date	Weekly_Sales	IsHoliday
1	1	1	05/02/2010	24924.5	false
2	1	1	12/02/2010	46039.49	true
3	1	1	19/02/2010	41595.55	false
4	1	1	26/02/2010	19403.54	false
5	1	1	05/03/2010	21827.9	false
6	1	1	12/03/2010	21043.39	false
7	1	1	19/03/2010	22136.64	false

stock

	Store	Type	Size
1	1	A	151315
2	2	A	202307
3	3	B	37392
4	4	A	205863
5	5	B	34875
6	6	A	202505
7	7	B	70713

features

	Store	Date	Temperature	Fuel_Price	MarkDown1	MarkDown2	MarkDown3	MarkDown4	MarkDown5	CPI	Unemployment	IsHoliday
1	1	05/02/2010	42.31	2.572	NA	NA	NA	NA	NA	211.0963582	8.106	false
2	1	12/02/2010	38.51	2.548	NA	NA	NA	NA	NA	211.2421698	8.106	true
3	1	19/02/2010	39.93	2.514	NA	NA	NA	NA	NA	211.2891429	8.106	false
4	1	26/02/2010	46.63	2.561	NA	NA	NA	NA	NA	211.3196429	8.106	false
5	1	05/03/2010	46.5	2.625	NA	NA	NA	NA	NA	211.3501429	8.106	false
6	1	12/03/2010	57.79	2.667	NA	NA	NA	NA	NA	211.3806429	8.106	false
7	1	19/03/2010	54.58	2.72	NA	NA	NA	NA	NA	211.215635	8.106	false

Implement a function to calculate average sales per store kind for weeks with holidays

The following dataset provides information on weekly sales from stores of a retail chain

sales

	Store	Dept	Date	Weekly_Sales	IsHoliday
1	1	1	05/02/2010	24924.5	false
2	1	1	12/02/2010	46039.49	true
3	1	1	19/02/2010	41595.55	false
4	1	1	26/02/2010	19403.54	false
5	1	1	05/03/2010	21827.9	false
6	1	1	12/03/2010	21043.39	false
7	1	1	19/03/2010	22136.64	false

store

	Store	Type	Size
1	1	A	151315
2	2	A	202307
3	3	B	37392
4	4	A	205863
5	5	B	34875
6	6	A	202505
7	7	B	70713

You are working in a team that is responsible for developing aggregate tables based on these two tables.
The aggregate information will be used as features for analytics.
One such request came recently that asks for the following:

Create a table that periodically updates average sales per store type.