DA5020 - Homework 5: Dates and Times

2019-09-21

Continue working with Farmers Market data from last week.

This week's assignment is not only about dates and times, but also what you learnt from past weeks: data transformation, strings, and more.

You may also need to go through a review on R control statesments since they will come handy in solving some of the problems.

Questions

- 1. (10 points) Add a new column Season1Days that contains the number of days a market is opened per week (for the dates it is open).
- 2. (10 points) Add a new column WeekendOpen indicating whether a market opens during weekends in Season1.
- 3. (20 points) Find out which markets close before 6PM, and which open only for fewer than 4 hours a day. For simplicity, consider only Season1Time. For markets with different open hours across a week, use the average length of open hours for the days they actually open.
- 4. (40 Points) The seasons are not standardized and would make analysis difficult. Create four new columns for four seasons (Spring, Summer, Fall, Winter), indicating whether a market is available in that season. Also, create two additional columns HalfYear and YearRound to identify those who open across seasons. Define "half year" and "year round" on your own terms, but explain them before you write the code (or as comments in your code). (Hint: you may want to create even more auxiliary columns, Season1BeginDate and Season1EndDate for example.)
- 5. (20 points) *Open question*: explore the new variables you just created. Aggregate them at different geographic levels, or some other categorical variable. What can you discover?

Submission

You need to submit an .Rmd extension file as well as the generated pdf file. Be sure to state all the assumptions and give explanations as comments in the .Rmd file wherever needed to help us assess your submission. Please name the submission file LAST_FirstInitial_1.Rmd for example for John Smith's 1st assignment, the file should be named Smith_J_1.Rmd.