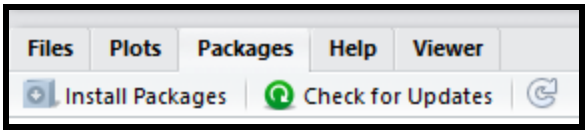
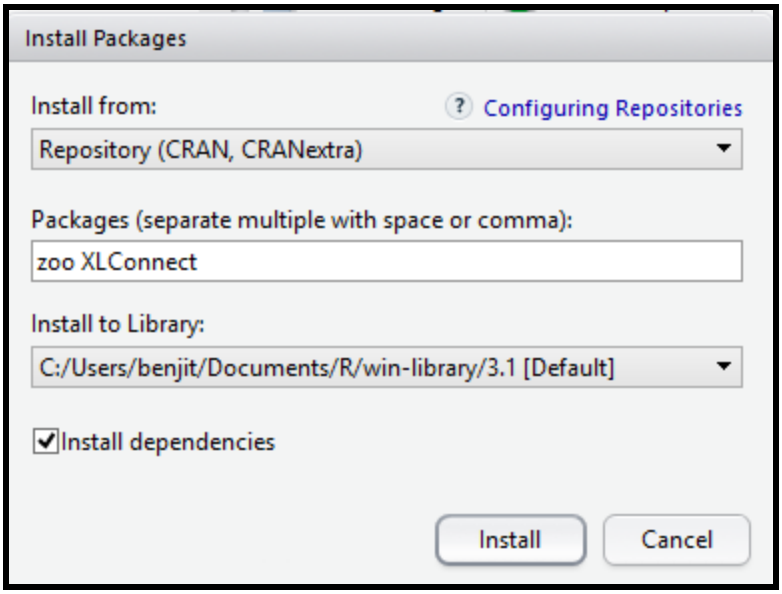
CleaningDataInR

Exercise 2: Reading Excel Files

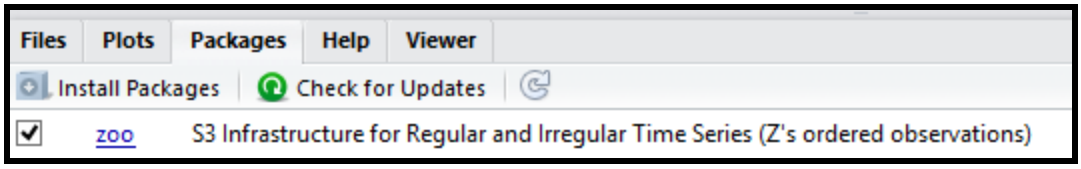
1. Open up R Studio and ex2.R

2. This exercise requires the use of the following R packages
   - XLConnect
     - to read excel file into R
     - ***install.packages('XLConnect')***
   - Zoo
     - To convert time into Year Quarter
     - ***install.packages('zoo')***
   - Alternatively you can use Rstudio to install

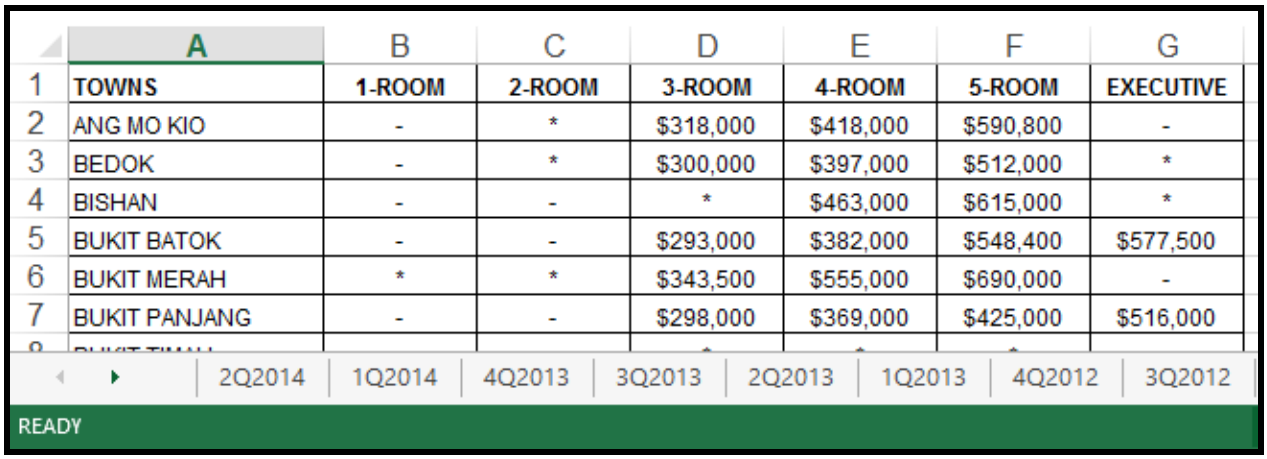




3. Load the library into R
   - ***library(XLConnect)***
   - ***library(Zoo)***
   - You can also use Rstudio to load the library by clicking on the checkbox



4. We need to read in the workbook into R and assign it to a variable
    - ***wb = loadWorkbook("C:\\Users\\benjit\\Google Drive\\CleaningDataWithR\\ex2\\MedianResalePrices.xls")***
5. The excel file has 29 sheets, we will need to read each sheet to extract the record
    - Please view the code with the comments on **ex2.R**



	A	B	C	D	E	F	G
1	TOWNS	1-ROOM	2-ROOM	3-ROOM	4-ROOM	5-ROOM	EXECUTIVE
2	ANG MO KIO	-	*	$318,000	$418,000	$590,800	-
3	BEDOK	-	*	$300,000	$397,000	$512,000	*
4	BISHAN	-	-	*	$463,000	$615,000	*
5	BUKIT BATOK	-	-	$293,000	$382,000	$548,400	$577,500
6	BUKIT MERAH	*	*	$343,500	$555,000	$690,000	-
7	BUKIT PANJANG	-	-	$298,000	$369,000	$425,000	$516,000

6. Lets create a time series for the resale price in Ang Mo Kio
    - We start from Year 2007 and Quarter 2 based on the spreadsheet
    - The frequency will be 4 since the data is quarterly data
    - ***AMK_TS = ts(data=AngMoKio,start = c(2007,2), frequency = 4)***
7. We convert the time series to a data frame as we want 2 columns to be written out, that is the date and the prices
    - ***AMK_DF = data.frame(YearQuarter=as.yearqtr(time(AMK_TS),format = "%Y-%m-%d"),AMK5RM=AMK_TS)***
8. Now will can write out as csv
    - ***write.csv(AMK_DF,"C:\\Users\\benjit\\Google Drive\\CleaningDataWithR\\ex2\\amkresalehousingdata.csv",row.names=FALSE)***