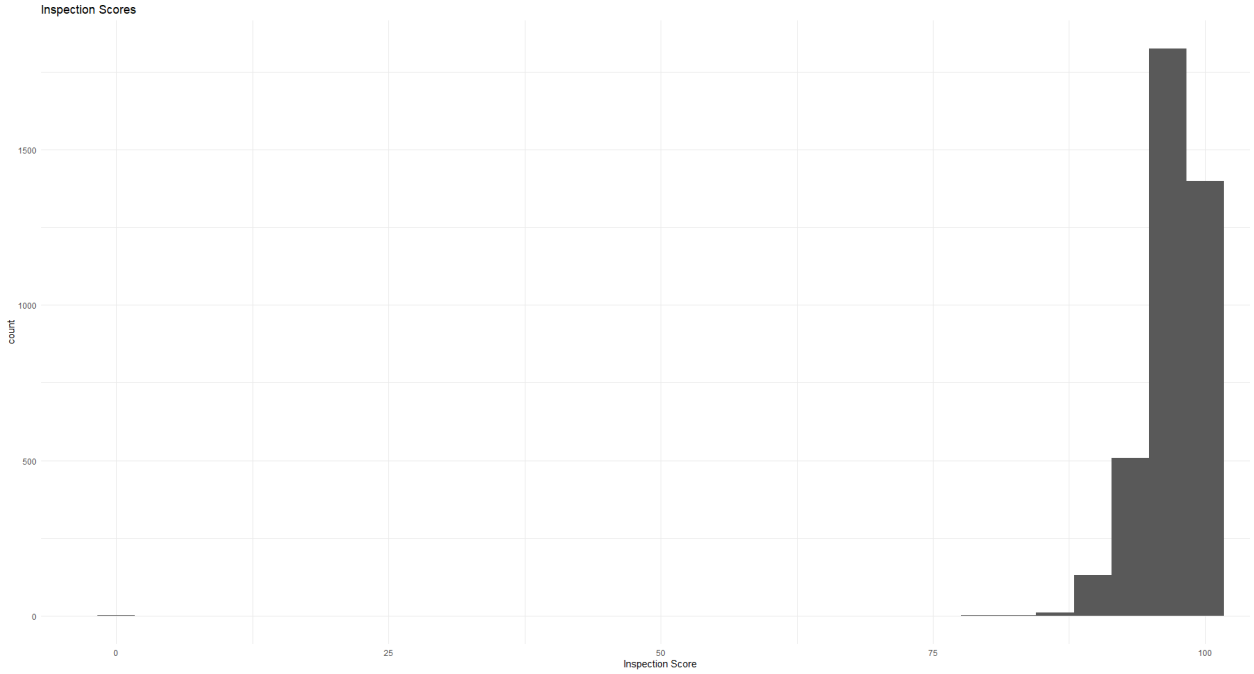
Alex Sang

Homework 2 Write-up

1.

Inspection scores are heavily left skewed as shown in the GGplot.



2.

The older facilities have a slightly higher average inspection score than newer facilities. I created two new tables, one with older open dates and one with newer open dates, and found that the mean score for older restaurants was higher than newer restaurants.

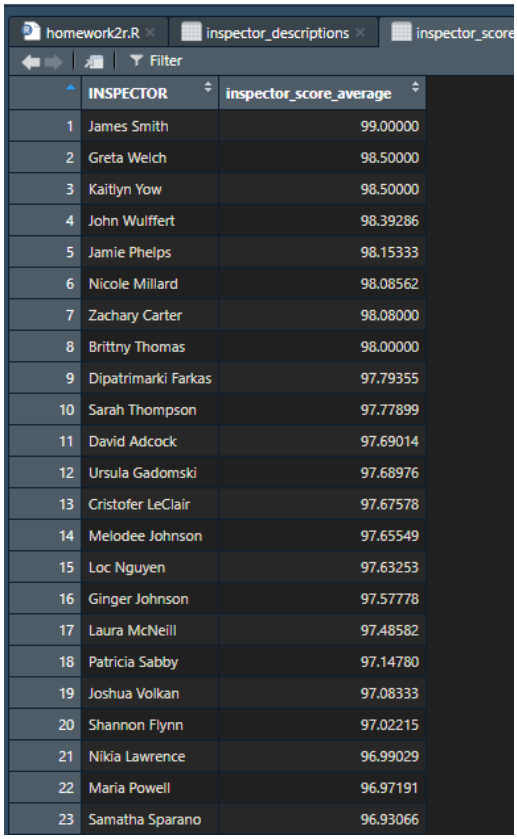
3.

Inspection scores do vary by city. After re-coding the cities and updating the table, the mean score was calculated, ordered and shown in table “city_score_mean.”

	newcity	average_score
1	NEW HILL	99.75000
2	RESEARCH TRIANGLE PARK	98.75000
3	HOLLY SPRINGS	98.35514
4	APEX	97.57568
5	CARY	97.55323
6	FUQUAY VARINA	97.34649
7	WAKE FOREST	97.30867
8	RALEIGH	97.12375
9	ROLESVILLE	97.08333
10	NORTH CAROLINA	97.00000
11	MORRISVILLE	96.75862
12	WILLOW SPRING	96.50000
13	GARNER	96.33835
14	KNIGHTDALE	96.15432
15	CLAYTON	96.12500
16	WENDELL	95.48571
17	ZEBULON	94.88000
18	ANGIER	94.50000
19	*NA*	94.15034

4.

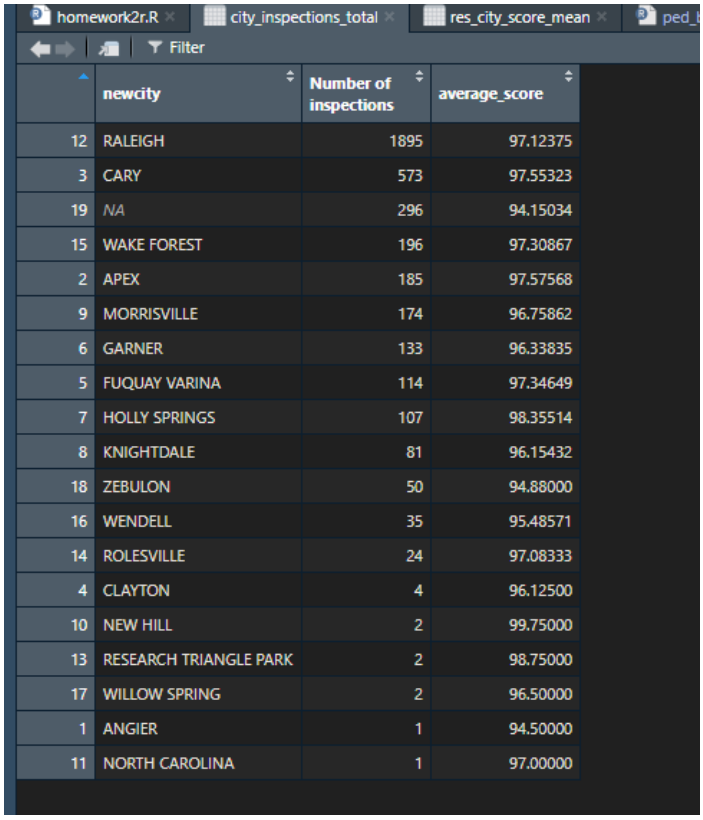
The inspection scores do vary by inspector and table “inspector_score” reveals the ordered average inspection score given by each inspector. I also included how many inspectors included descriptions by counting descriptions per inspector through the table “inspector_descriptions.”



	INSPECTOR	inspector_score_average
1	James Smith	99.00000
2	Greta Welch	98.50000
3	Kaitlyn Yow	98.50000
4	John Wulffert	98.39286
5	Jamie Phelps	98.15333
6	Nicole Millard	98.08562
7	Zachary Carter	98.08000
8	Brittny Thomas	98.00000
9	Dipatrimarki Farkas	97.79355
10	Sarah Thompson	97.77899
11	David Adcock	97.69014
12	Ursula Gadomski	97.68976
13	Cristofer LeClair	97.67578
14	Melodee Johnson	97.65549
15	Loc Nguyen	97.63253
16	Ginger Johnson	97.57778
17	Laura McNeill	97.48582
18	Patricia Sabby	97.14780
19	Joshua Volkan	97.08333
20	Shannon Flynn	97.02215
21	Nikia Lawrence	96.99029
22	Maria Powell	96.97191
23	Samatha Sparano	96.93066

5.

Yes some extreme results from the previous questions are due to small sample sizes. Table “city_inspections_total” orders the cities by the number of inspections and it can be seen that the sample sizes greatly differ.



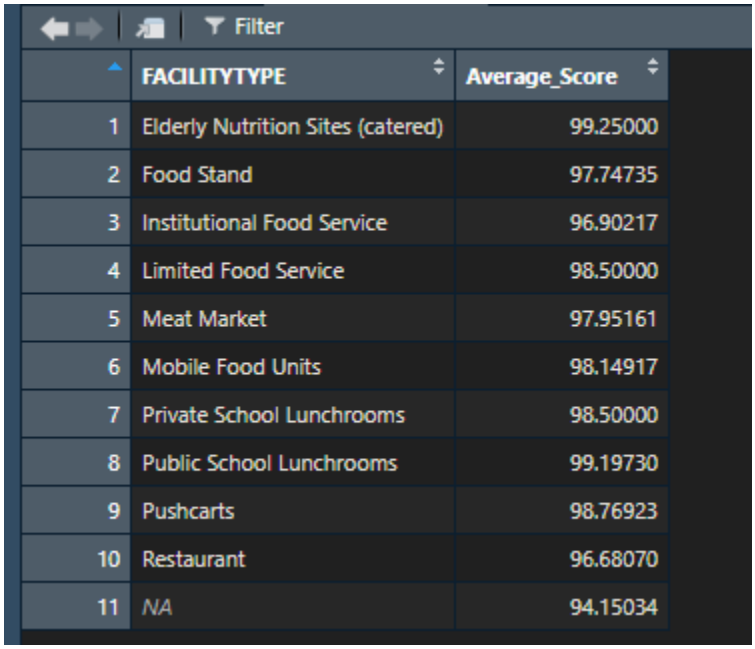
homework2r.R | city_inspections_total | res_city_score_mean | ped_l

Filter

	newcity	Number of inspections	average_score
12	RALEIGH	1895	97.12375
3	CARY	573	97.55323
19	*NA*	296	94.15034
15	WAKE FOREST	196	97.30867
2	APEX	185	97.57568
9	MORRISVILLE	174	96.75862
6	GARNER	133	96.33835
5	FUQUAY VARINA	114	97.34649
7	HOLLY SPRINGS	107	98.35514
8	KNIGHTDALE	81	96.15432
18	ZEBULON	50	94.88000
16	WENDELL	35	95.48571
14	ROLESVILLE	24	97.08333
4	CLAYTON	4	96.12500
10	NEW HILL	2	99.75000
13	RESEARCH TRIANGLE PARK	2	98.75000
17	WILLOW SPRING	2	96.50000
1	ANGIER	1	94.50000
11	NORTH CAROLINA	1	97.00000

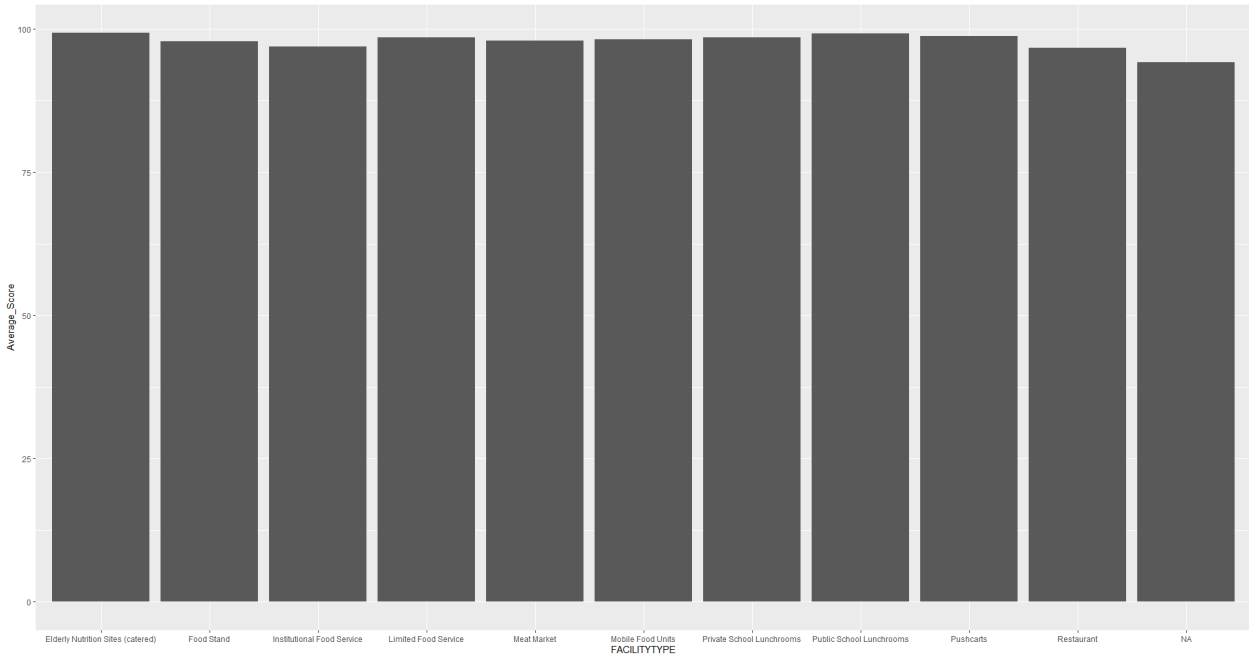
6.

No, table “facility_mean” shows that restaurants rank at the bottom for average inspection scores and the GGplot shows the same results.



Filter

	FACILITYTYPE	Average_Score
1	Elderly Nutrition Sites (catered)	99.25000
2	Food Stand	97.74735
3	Institutional Food Service	96.90217
4	Limited Food Service	98.50000
5	Meat Market	97.95161
6	Mobile Food Units	98.14917
7	Private School Lunchrooms	98.50000
8	Public School Lunchrooms	99.19730
9	Pushcarts	98.76923
10	Restaurant	96.68070
11	*NA*	94.15034



7-2

Newer restaurants have a better inspection score average

7-3

Inspection scores vary by city shown in “res_city_score_mean”

7-4

Inspection scores vary by inspector shown in “res_inspector_score”

7-5

Sample sizes vary by city and explain the results for mean inspection scores shown in table “res_city_inspections_total”