Data Normalization

Team Members :-

Aakash Rajawat- 002764127

Diya Baldota- 002747966

Hardik Sodhani- 002770306

Rucha Chotalia- 002711888

Table: Housing Data

The housing data was not in 1NF as it had multiple data in single column

Solution: Data was normalized using python script to bring in 1NF form

https://colab.research.google.com/drive/1CiiHFNYYM0Xm1kVO_wpWN4JUUlL3zaz7#scrollTo=wYWrxaj RF3yD

Data:-

Before Normalizing

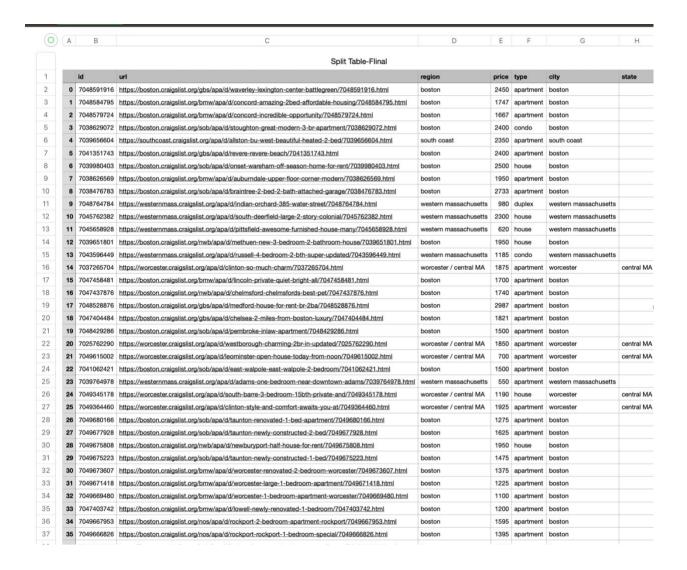
	Α	В	С	D	Е	F
3536	7050927500	https://worcester.craigslist.org/apa/d/bellingham-the-charles-bellingham-in/7050927500.html	worcester / central MA	1769	apartment	
3537	7050927311	https://worcester.craigslist.org/apa/d/ashland-wonderful-2bedroom-nice-and/7050927311.html	worcester / central MA	1350	house	
3538	7048988994	https://worcester.craigslist.org/apa/d/westborough-private-balcony-rec-room/7048988994.html	worcester / central MA	1630	apartment	
3539	7040054504	https://worcester.craigslist.org/apa/d/worcester-15-forkey-ave-3-bd-2-ba-for/7040054504.html	worcester / central MA	850	apartment	
3540	7040048768	https://worcester.craigslist.org/apa/d/worcester-7-creswell-rd-2-bd-2-ba-for/7040048768.html	worcester / central MA	800	apartment	
3541	7034478068	https://worcester.craigslist.org/apa/d/worcester-1-walter-st-3-bd-2-ba-for/7034478068.html	worcester / central MA	900	apartment	
3542	7034475049	https://worcester.craigslist.org/apa/d/worcester-19-passway-3-2-bd-1-ba-for/7034475049.html	worcester / central MA	750	apartment	
3543	7045031423	https://worcester.craigslist.org/apa/d/auburn-auburn-beautiful-includes/7045031423.html	worcester / central MA	1000	house	
3544	7041334168	https://worcester.craigslist.org/apa/d/bellingham-apartments-available-in/7041334168.html	worcester / central MA	1935	apartment	
545	7048603216	https://worcester.craigslist.org/apa/d/grafton-housing-available-for-immediate/7048603216. html the state of the state o	worcester / central MA	1300	apartment	
546	7048600678	https://worcester.craigslist.org/apa/d/boylston-great-place-available-today/7048600678.html	worcester / central MA	1100	townhouse	
547	7050712040	https://worcester.craigslist.org/apa/d/worcester-spacey-super-345-bed-room/7050712040.html	worcester / central MA	640	house	
548	7050871054	https://worcester.craigslist.org/apa/d/worcester-leasing-now-hair-salon-elder/7050871054.html	worcester / central MA	1500	apartment	
549	7050863243	https://worcester.craigslist.org/apa/d/belling ham-apartments-in-belling ham-24/7050863243. https://www.apartments-in-belling ham-24/705086	worcester / central MA	1979	apartment	
550	7050798045	https://worcester.craigslist.org/apa/d/worcester-wonderful-apartment-for-rent/7050798045.htm.	worcester / central MA	1680	apartment	
551	7050797641	https://worcester.craigslist.org/apa/d/worcester-2-bedrooms-2-bathrooms-pet/7050797641.htm.	worcester / central MA	1245	apartment	
552	7048300281	https://worcester.craigslist.org/apa/d/maynard-move-in-special-in-unit-washer/7048300281.htm.	worcester / central MA	3019	apartment	
553	7050853060	https://worcester.craigslist.org/apa/d/maynard-ask-about-our-move-in-special/7050853060.htm	worcester / central MA	2369	apartment	
554	7050842602	https://worcester.craigslist.org/apa/d/marlborough-rt-20-huge-modern-2-bed/7050842602.html.	worcester / central MA	1475	apartment	
555	7050842533	https://worcester.craigslist.org/apa/d/marlborough-rt-20-huge-modern-2-bed/7050842533.html	worcester / central MA	1475	apartment	
556	7050838203	https://worcester.craigslist.org/apa/d/shrewsbury-immediate-move-in-2-br/7050838203.html	worcester / central MA	1556	apartment	
557	7035042700	https://worcester.craigslist.org/apa/d/worcester-46-wrentham-rd-bd-2-ba-for/7035042700.html	worcester / central MA	900	apartment	
558	7035044443	https://worcester.craigslist.org/apa/d/worcester-amazing-view-clara-st-3-bd-2/7035044443.html	worcester / central MA	1000	apartment	
559	7048297852	https://worcester.craigslist.org/ana/d/hellingham-your-new-anartment-home-in/70/8297852.htm	worrester / central MA	1940	anartment	

After Normalizing in 1NF:-

	id	url	region	price	type	city	state
0	7048591916	https://boston.craigslist.org/gbs/apa/d/waverl	boston	2450	apartment	boston	None
1	7048584795	https://boston.craigslist.org/bmw/apa/d/concor	boston	1747	apartment	boston	None
2	7048579724	https://boston.craigslist.org/bmw/apa/d/concor	boston	1667	apartment	boston	None
3	7038629072	https://boston.craigslist.org/sob/apa/d/stough	boston	2400	condo	boston	None
4	7039656604	https://southcoast.craigslist.org/apa/d/allsto	south coast	2350	apartment	south coast	None
4921	7024971933	https://worcester.craigslist.org/apa/d/fitchbu	worcester / central MA	1480	apartment	worcester	central MA
4922	7024959372	https://worcester.craigslist.org/apa/d/hudson	worcester / central MA	2155	apartment	worcester	central MA
4923	7024940732	https://worcester.craigslist.org/apa/d/worcest	worcester / central MA	1550	apartment	worcester	central MA
4924	7024932794	https://worcester.craigslist.org/apa/d/worcest	worcester / central MA	1810	apartment	worcester	central MA
4925	7024928221	https://worcester.craigslist.org/apa/d/berlin	worcester / central MA	1	apartment	worcester	central MA

3rd Normal Form:

- For the table to be in 3rd Normal Form, it must satisfy the conditions of 1st Normal Form and 2nd Normal Form.
- The table is now satisfying the 2nd Normal Form.
- There are no partial and transitive dependencies in the table. So the table is now in 3rd Normal Form.



2nd Normal Form:

- · For the table to be in 2nd Normal Form, it must satisfy the conditions of 1st Normal Form.
- The table has now been converted into 2nd Normal Form
- · This table does not contain partial dependency.

As this table, fulfils all the above criteria, it is said to be in 2nd Normal Form.

	id	urt	region	price	type	city	state
0	7048591916	https://boston.craigslist.org/gbs/apa/d/waverley-lexington-center-battlegreen/7048591916.html	boston	2450	apartment	boston	
1	7048584795	https://boston.craigslist.org/bmw/apa/d/concord-amazing-2bed-affordable-housing/7048584795.html	boston	1747	apartment	boston	
2	7048579724	https://boston.craigslist.org/bmw/apa/d/concord-incredible-opportunity/7048579724.html	boston	1667	apartment	boston	
3	7038629072	https://boston.craigslist.org/sob/apa/d/stoughton-great-modern-3-br-apartment/7038629072.html	boston	2400	condo	boston	
4	7039656604	https://southcoast.craigslist.org/apa/d/allston-bu-west-beautiful-heated-2-bed/7039656604.html	south coast	2350	apartment	south coast	
5	7041351743	https://boston.craigslist.org/gbs/apa/d/revere-revere-beach/7041351743.html	boston	2400	apartment	boston	
6	7039980403	https://boston.craigslist.org/sob/apa/d/onset-wareham-off-season-home-for-rent/7039980403.html	boston	2500	house	boston	
7	7038626569	https://boston.craigslist.org/bmw/apa/d/auburndale-upper-floor-corner-modern/7038626569.html	boston	1950	apartment	boston	
8	7038476783	https://boston.craigslist.org/sob/apa/d/braintree-2-bed-2-bath-attached-garage/7038476783.html	boston	2733	apartment	boston	
9	7048764784	https://westernmass.craigslist.org/apa/d/indian-orchard-385-water-street/7048764784.html	western massachusetts	980	duplex	western massachusetts	
0	7045762382	https://westernmass.craigslist.org/apa/d/south-deerfield-large-2-story-colonial/7045762382.html	western massachusetts	2300	house	western massachusetts	
1	7045658928	https://westernmass.craigslist.org/apa/d/pittsfield-awesome-furnished-house-many/7045658928.html	western massachusetts	620	house	western massachusetts	
2	7039651801	https://boston.craigslist.org/nwb/apa/d/methuen-new-3-bedroom-2-bathroom-house/7039651801.html	boston	1950	house	boston	
3	7043596449	https://westernmass.craigslist.org/apa/d/russell-4-bedroom-2-bth-super-updated/7043596449.html	western massachusetts	1185	condo	western massachusetts	
4	7037265704	https://worcester.craigslist.org/apa/d/clinton-so-much-charm/7037265704.html	worcester / central MA	1875	apartment	worcester	central MA
5	7047458481	https://boston.craigslist.org/bmw/apa/d/lincoln-private-quiet-bright-all/7047458481.html	boston	1700	apartment	boston	
6	7047437876	https://boston.craigslist.org/nwb/apa/d/chelmsford-chelmsfords-best-pet/7047437876.html	boston	1740	apartment	boston	
7	7048528876	https://boston.craigslist.org/gbs/apa/d/medford-house-for-rent-br-2ba/7048528876.html	boston	2987	apartment	boston	
8	7047404484	https://boston.craigslist.org/gbs/apa/d/chelsea-2-miles-from-boston-luxury/7047404484.html	boston	1821	apartment	boston	
9	7048429286	https://boston.craigslist.org/sob/apa/d/pembroke-inlaw-apartment/7048429286.html	boston	1500	apartment	boston	
0	7025762290	https://worcester.craigslist.org/apa/d/westborough-charming-2br-in-updated/7025762290.html	worcester / central MA	1850	apartment	worcester	central MA
1	7049615002	https://worcester.craigslist.org/apa/d/leominster-open-house-today-from-noon/7049615002.html	worcester / central MA	700	apartment	worcester	central MA
2	7041062421	https://boston.craigslist.org/sob/apa/d/east-walpole-east-walpole-2-bedroom/7041062421.html	boston	1500	apartment	boston	
3	7039764978	https://westernmass.craigslist.org/apa/d/adams-one-bedroom-near-downtown-adams/7039764978.html	western massachusetts	550	apartment	western massachusetts	
4	7049345178	https://worcester.craigslist.org/apa/d/south-barre-3-bedroom-15bth-private-and/7049345178.html	worcester / central MA	1190	house	worcester	central MA
5	7049364460	https://worcester.craigslist.org/apa/d/clinton-style-and-comfort-awaits-you-at/7049364460.html	worcester / central MA	1925	apartment	worcester	central MA
6	7049680166	https://boston.craigslist.org/sob/apa/d/taunton-renovated-1-bed-apartment/7049680166.html	boston	1275	apartment	boston	
7	7049677928	https://boston.craigslist.org/sob/apa/d/taunton-newly-constructed-2-bed/7049677928.html	boston	1625	apartment	boston	
8	7049675808	https://boston.craigslist.org/nwb/apa/d/newburyport-half-house-for-rent/7049675808.html	boston	1950	house	boston	
9	7049675223	https://boston.craigslist.org/sob/apa/d/taunton-newly-constructed-1-bed/7049675223.html	boston	1475	apartment	boston	
0	7049673607	https://boston.craigslist.org/bmw/apa/d/worcester-renovated-2-bedroom-worcester/7049673607.html	boston	1375	apartment	boston	
1	7049671418	https://boston.craigslist.org/bmw/apa/d/worcester-large-1-bedroom-apartment/7049671418.html	boston	1225	apartment	boston	
2	7049669480	https://boston.craigslist.org/bmw/apa/d/worcester-1-bedroom-apartment-worcester/7049669480.html	boston	1100	apartment	boston	
3	7047403742	https://boston.craigslist.org/bmw/apa/d/lowell-newly-renovated-1-bedroom/7047403742.html	boston	1200	apartment	boston	
4	7049667953	https://boston.craigslist.org/nos/apa/d/rockport-2-bedroom-apartment-rockport/7049667953.html	boston	1595	apartment	boston	
5	7049666826	https://boston.craigslist.org/nos/apa/d/rockport-rockport-1-bedroom-special/7049666826.html	boston	1395	apartment	boston	
						100	

VIEW SQL queries:

1. CREATE VIEW beds AS(

SELECT dmddapt1.id, dmddapt2.beds, dmddapt2.sqfeet FROM dmddapt1 INNER JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id);

2. CREATE VIEW laundry AS(

SELECT dmddapt1.id, dmddapt2.beds, dmddapt2.laundry_options FROM dmddapt1 LEFT JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id);

3. CREATE VIEW region AS(

SELECT dmddapt1.id, dmddapt2.region, dmddapt2.sqfeet FROM dmddapt2 RIGHT JOIN dmddapt1 ON dmddapt1.id=dmddapt2.id);

4. CREATE VIEW sqfeet AS(

SELECT dmddapt1.id, dmddapt2.region, dmddapt2.sqfeet FROM dmddapt2 INNER JOIN dmddapt1 ON dmddapt1.id=dmddapt2.id ORDER BY dmddapt2.sqfeet DESC);

5. CREATE VIEW area AS(

SELECT dmddapt1.type, dmddapt2.sqfeet FROM dmddapt2 INNER JOIN dmddapt1 ON dmddapt1.id=dmddapt2.id);

6. CREATE VIEW price AS(

SELECT AVG(dmddapt1.price) FROM dmddapt2 INNER JOIN dmddapt1 ON dmddapt1.id=dmddapt2.id);

7. CREATE VIEW state AS(

SELECT state, count(image_url) as count FROM dmddapt4 GROUP BY state);

8. CREATE VIEW id AS(

SELECT COUNT(id) as count, region FROM dmddapt2 GROUP BY region);

9. CREATE VIEW carpet AS(

SELECT dmddapt1.id, dmddapt2.beds, dmddapt2.sqfeet FROM dmddapt1 INNER JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.laundry_options like '%laundry in bldg%' or dmddapt2.laundry_options like '%laundry on site%');

10. CREATE VIEW bhk AS(

SELECT dmddapt1.id, dmddapt1.price, dmddapt1.type, dmddapt2.region FROM dmddapt1 INNER JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.region = 'Boston' and dmddapt1.type like '%apartment%');

11. CREATE VIEW apperance AS(

SELECT dmddapt1.id, dmddapt2.region, dmddapt2.sqfeet FROM dmddapt2 RIGHT JOIN dmddapt1 ON dmddapt1.id=dmddapt2.id);

12. CREATE VIEW apartment AS(

SELECT * FROM dmddapt1 JOIN dmddapt4 ON dmddapt1.id=dmddapt4.id where dmddapt1.type = 'apartment' and dmddapt4.image_url IS NOT NULL);

13. CREATE VIEW boston AS(

SELECT dmddapt1.id, dmddapt1.price, dmddapt1.type, dmddapt2.region FROM dmddapt1 JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.region = 'Boston' and dmddapt1.type like '%apartment%');

14. CREATE VIEW morethanprice AS(

SELECT dmddapt1.id, dmddapt1.price, dmddapt2.region FROM dmddapt1 JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.sqfeet > 700 and dmddapt1.type like '%apartment%' ORDER BY dmddapt1.price DESC);

15. CREATE VIEW unit AS(

SELECT dmddapt1.id, dmddapt1.price, dmddapt2.beds,dmddapt2.laundry_options FROM dmddapt1 JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.laundry_options='w/d in unit');

16. CREATE VIEW bedsthan AS(

SELECT dmddapt1.id, dmddapt1.price, dmddapt2.beds,dmddapt2.beds FROM dmddapt1 JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.beds>1);

17. CREATE VIEW regioning AS(

SELECT DISTINCT dmddapt2.region, dmddapt2.sqfeet FROM dmddapt2 RIGHT JOIN dmddapt1 ON dmddapt1.id=dmddapt2.id);

18. CREATE VIEW petting AS(

SELECT dmddapt1.id, dmddapt3.cats_allowed, dmddapt3.cats_allowed FROM dmddapt3 RIGHT JOIN dmddapt1 ON dmddapt1.id=dmddapt3.id);

19. CREATE VIEW bedss AS(

SELECT dmddapt1.id, dmddapt1.price, dmddapt2.beds,dmddapt2.beds FROM dmddapt1 JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id where dmddapt2.beds>1);

20. CREATE VIEW idssss AS(

SELECT dmddapt1.id, dmddapt2.beds, dmddapt2.laundry_options FROM dmddapt1 LEFT JOIN dmddapt2 ON dmddapt1.id=dmddapt2.id);

ER diagram:

