Appendix of 'The effects of mismatched train and test data cleaning pipelines on regression models: lessons for practice'

Cleaning Setups - Airbnb Data

Setup number	mv_repair	$outlier_detection$	$outlier_repair$	${\bf duplicate_repair}$
0	delete	none	NA	NA
1	delete	none	NA	key_val
2	delete	SD	mean	NA
3	delete	SD	mean	key_val
4	delete	SD	median	NA
5	delete	SD	median	key_val
6	delete	SD	mode	NA
7	delete	SD	mode	key_val
8	delete	IQR	mean	NA
9	delete	IQR	mean	key_val
10	delete	IQR	median	NA
11	delete	IQR	median	key_val
12	delete	IQR	mode	NA
13	delete	IQR	mode	key_val
14	mean-mode	none	NA	NA
15	mean-mode	none	NA	key_val
16	mean-mode	SD	mean	NA
17	mean-mode	SD	mean	key_val
18	mean-mode	SD	median	NA
19	mean-mode	SD	median	key_val
20	mean-mode	SD	mode	NA
21	mean-mode	SD	mode	key_val
22	mean-mode	IQR	mean	NA
23	mean-mode	IQR	mean	key_val
24	mean-mode	IQR	median	NA
25	mean-mode	IQR	median	key_val
26	mean-mode	IQR	mode	NA
27	mean-mode	IQR	mode	key_val
28	median-mode	none	NA	NA
29	median-mode	none	NA	key_val
30	median-mode	SD	mean	NA
31	median-mode	SD	mean	key_val
32	median-mode	SD	median	NA
33	median-mode	SD	median	key_val
34	median-mode	SD	mode	NA
35	median-mode	SD	mode	key_val

37 median-mode IQR mean key_val 38 median-mode IQR median NA 39 median-mode IQR median key_val 40 median-mode IQR mode key_val 40 median-mode IQR mode key_val 40 median-mode IQR mode key_val 41 median-mode IQR mode key_val 42 mode-mode SD mean NA 43 mode-mode SD mean key_val 45 mode-mode SD median key_val 46 mode-mode SD median key_val 47 mode-mode SD mode key_val 48 mode-mode IQR mean key_val 50 mode-mode IQR median NA 51 mode-mode IQR median key_val 53	36	median-mode	IQR	mean	NA
39 median-mode IQR median key_val 40 median-mode IQR mode NA 41 median-mode IQR mode key_val 41 median-mode IQR mode key_val 42 mode-mode none NA key_val 43 mode-mode SD mean key_val 45 mode-mode SD median NA 45 mode-mode SD median key_val 46 mode-mode SD median key_val 48 mode-mode SD mode NA 49 mode-mode IQR mean NA 50 mode-mode IQR mean key_val 51 mode-mode IQR median key_val 53 mode-mode IQR median key_val 54 mode-mode IQR median key_val 55 <	37	median-mode	IQR	mean	key_val
40 median-mode IQR mode key_val 41 median-mode IQR mode key_val 42 mode-mode IQR mode key_val 42 mode-mode none NA NA 43 mode-mode SD mean NA 44 mode-mode SD median NA 45 mode-mode SD median NA 46 mode-mode SD median NA 46 mode-mode SD median NA 48 mode-mode SD mode key_val 49 mode-mode IQR mean NA 50 mode-mode IQR median NA 51 mode-mode IQR median key_val 52 mode-mode IQR mode key_val 54 mode-mode IQR mode key_val 55 mode-mode <	38	median-mode	IQR	median	NA
41 median-mode IQR mode key_val 42 mode-mode none NA NA 43 mode-mode none NA key_val 44 mode-mode SD mean key_val 45 mode-mode SD median NA 46 mode-mode SD median NA 47 mode-mode SD median key_val 48 mode-mode SD mode key_val 49 mode-mode SD mode key_val 50 mode-mode IQR mean NA 51 mode-mode IQR median key_val 52 mode-mode IQR median key_val 53 mode-mode IQR mode NA 54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dummy	39	median-mode	IQR	median	key_val
42 mode-mode none NA NA 43 mode-mode none NA key_val 44 mode-mode SD mean NA 45 mode-mode SD median key_val 46 mode-mode SD median key_val 47 mode-mode SD median key_val 48 mode-mode SD mode NA 48 mode-mode SD mode NA 49 mode-mode IQR mean NA 50 mode-mode IQR mean key_val 51 mode-mode IQR median key_val 53 mode-mode IQR median key_val 54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy <t< td=""><td>40</td><td>median-mode</td><td>IQR</td><td>mode</td><td>NA</td></t<>	40	median-mode	IQR	mode	NA
43 mode-mode No key_val 44 mode-mode SD mean NA 45 mode-mode SD mean key_val 46 mode-mode SD median NA 47 mode-mode SD median key_val 48 mode-mode SD mode NA 49 mode-mode SD mode key_val 50 mode-mode IQR mean NA 51 mode-mode IQR median NA 51 mode-mode IQR median key_val 52 mode-mode IQR median key_val 53 mode-mode IQR mode NA 54 mode-mode IQR mode key_val 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy SD <td< td=""><td>41</td><td>median-mode</td><td>IQR</td><td>mode</td><td>key_val</td></td<>	41	median-mode	IQR	mode	key_val
44 mode-mode SD mean NA 45 mode-mode SD mean key_val 46 mode-mode SD median NA 47 mode-mode SD median NA 48 mode-mode SD mode NA 49 mode-mode IQR mean NA 50 mode-mode IQR mean NA 51 mode-mode IQR median NA 52 mode-mode IQR median key_val 53 mode-mode IQR median key_val 54 mode-mode IQR median key_val 55 mode-mode IQR median key_val 56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD median NA 60 mean-dummy <td< td=""><td>42</td><td>$\operatorname{mode-mode}$</td><td>none</td><td>NA</td><td>NA</td></td<>	42	$\operatorname{mode-mode}$	none	NA	NA
45 mode-mode SD mean key_val 46 mode-mode SD median NA 47 mode-mode SD median key_val 48 mode-mode SD mode NA 48 mode-mode SD mode key_val 50 mode-mode IQR mean NA 50 mode-mode IQR mean NA 51 mode-mode IQR median NA 52 mode-mode IQR median key_val 53 mode-mode IQR median key_val 54 mode-mode IQR median key_val 55 mode-mode IQR median key_val 54 mode-mode IQR median key_val 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA NA 57	43	$\operatorname{mode-mode}$	none	NA	key_val
46 mode-mode SD median NA 47 mode-mode SD median key_val 48 mode-mode SD mode NA 49 mode-mode IQR mean NA 50 mode-mode IQR mean key_val 50 mode-mode IQR median NA 51 mode-mode IQR median NA 52 mode-mode IQR median key_val 53 mode-mode IQR mode NA 54 mode-mode IQR mode NA 55 mode-mode IQR mode NA 56 mean-dum none NA NA 57 mode-mode IQR mode key_val 58 mean-dummy none NA key_val 58 mean-dummy SD mean key_val 60 mean-dummy SD <td>44</td> <td>mode-mode</td> <td>SD</td> <td>mean</td> <td>NA</td>	44	mode-mode	SD	mean	NA
47 mode-mode SD median key_val 48 mode-mode SD mode NA 49 mode-mode IQR mean NA 50 mode-mode IQR mean NA 51 mode-mode IQR median NA 51 mode-mode IQR median key_val 52 mode-mode IQR median key_val 53 mode-mode IQR mode NA 54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dum NA NA NA 56 mean-dummy none NA NA 57 mean-dummy ND mean key_val 68 mean-dummy SD median key_val 60 mean-dummy SD mode key_val 63 mean-dummy I	45	mode-mode	SD	mean	key_val
48 mode-mode SD mode NA 49 mode-mode SD mode key_val 50 mode-mode IQR mean NA 51 mode-mode IQR median NA 52 mode-mode IQR median key_val 53 mode-mode IQR median key_val 54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 59 mean-dummy SD median NA 60 mean-dummy SD median key_val 61 mean-dummy SD mode NA 62 mean-dummy SD median NA 63 mean-dummy	46	mode-mode	SD	median	NA
49 mode-mode SD mode key_val 50 mode-mode IQR mean NA 51 mode-mode IQR mean key_val 52 mode-mode IQR median NA 53 mode-mode IQR median key_val 54 mode-mode IQR mode key_val 55 mode-mode IQR mode NA 56 mean-dummy none NA NA 57 mean-dummy none NA NA 58 mean-dummy SD mean key_val 59 mean-dummy SD median NA 60 mean-dummy SD median key_val 61 mean-dummy SD median key_val 62 mean-dummy SD mean key_val 63 mean-dummy IQR mean key_val 65 mean-dummy <td>47</td> <td>mode-mode</td> <td>SD</td> <td></td> <td>key_val</td>	47	mode-mode	SD		key_val
50 mode-mode IQR mean NA 51 mode-mode IQR mean key_val 52 mode-mode IQR median NA 53 mode-mode IQR median key_val 54 mode-mode IQR mode key_val 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 60 mean-dummy SD median NA 61 mean-dummy SD mode key_val 62 mean-dummy SD mode key_val 63 mean-dummy IQR mean NA 64 mean-dummy IQR median NA 65 mean-dummy	48	mode-mode			NA
51 mode-mode IQR mean key_val 52 mode-mode IQR median NA 53 mode-mode IQR median key_val 54 mode-mode IQR mode NA 55 mode-mode IQR mode NA 55 mode-mode IQR mode NA 56 mean-dummy none NA NA 57 mean-dummy none NA NA 58 mean-dummy SD mean NA 60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode key_val 63 mean-dummy IQR mean NA 64 mean-dummy IQR median NA 65 mean-dummy IQR median NA 66 mean-dummy IQR </td <td>49</td> <td>mode-mode</td> <td>SD</td> <td>mode</td> <td>key_val</td>	49	mode-mode	SD	mode	key_val
52 mode-mode IQR median NA 53 mode-mode IQR median key_val 54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy SD mean NA 58 mean-dummy SD mean NA 59 mean-dummy SD median NA 60 mean-dummy SD median NA 61 mean-dummy SD mode NA 62 mean-dummy SD mode key_val 63 mean-dummy IQR mean NA 64 mean-dummy IQR median NA 65 mean-dummy IQR median key_val 66 mean-dummy IQR median key_val 69 mean-dummy <	50	mode-mode	IQR	mean	NA
53 mode-mode IQR median key_val 54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 59 mean-dummy SD mean key_val 60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode NA 63 mean-dummy SD mode key_val 64 mean-dummy IQR mean NA 65 mean-dummy IQR median NA 66 mean-dummy IQR median NA 67 mean-dummy IQR mode NA 69 mean-dummy	51	mode-mode	IQR	mean	key_val
54 mode-mode IQR mode NA 55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 59 mean-dummy SD mean key_val 60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode key_val 63 mean-dummy SD mode key_val 64 mean-dummy IQR mean NA 65 mean-dummy IQR median NA 66 mean-dummy IQR median NA 67 mean-dummy IQR median key_val 68 mean-dummy IQR mode key_val 70 median-dummy<		mode-mode	IQR	median	NA
55 mode-mode IQR mode key_val 56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 59 mean-dummy SD median NA 60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode key_val 63 mean-dummy IQR mean NA 64 mean-dummy IQR mean key_val 65 mean-dummy IQR median key_val 66 mean-dummy IQR median key_val 69 mean-dummy IQR mode key_val 70 median-dummy none NA NA 71 median-dummy none NA key_val 72 medi	53	mode-mode	IQR	median	key_val
56 mean-dummy none NA NA 57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 59 mean-dummy SD median NA 60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode key_val 63 mean-dummy IQR mean NA 64 mean-dummy IQR mean key_val 65 mean-dummy IQR median key_val 66 mean-dummy IQR median key_val 68 mean-dummy IQR mode key_val 69 mean-dummy IQR mode key_val 70 median-dummy none NA NA 71 median-dummy none NA key_val 72 med	54	mode-mode	IQR	mode	NA
57 mean-dummy none NA key_val 58 mean-dummy SD mean NA 59 mean-dummy SD mean key_val 60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode NA 63 mean-dummy SD mode key_val 64 mean-dummy IQR mean NA 65 mean-dummy IQR median NA 66 mean-dummy IQR median NA 67 mean-dummy IQR median NA 68 mean-dummy IQR median NA 69 mean-dummy IQR mode NA 69 mean-dummy IQR mode NA 70 median-dummy none NA NA 71 median-dummy <	55	mode-mode	IQR	mode	key_val
58mean-dummySDmeanNA59mean-dummySDmeankey_val60mean-dummySDmedianNA61mean-dummySDmediankey_val62mean-dummySDmodeNA63mean-dummySDmodekey_val64mean-dummyIQRmeanNA65mean-dummyIQRmedianNA66mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNANA72median-dummySDmeanNA73median-dummySDmeankey_val74median-dummySDmedianNA75median-dummySDmedianNA76median-dummySDmodeNA77median-dummySDmodeNA78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmeankey_val	56	mean-dummy	none	NA	NA
59mean-dummySDmeankey_val60mean-dummySDmedianNA61mean-dummySDmediankey_val62mean-dummySDmodeNA63mean-dummySDmodekey_val64mean-dummyIQRmeanNA65mean-dummyIQRmedianNA66mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNANA72median-dummySDmeanNA73median-dummySDmedianNA75median-dummySDmedianNA75median-dummySDmodeNA76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmeankey_val	57	mean-dummy	none	NA	key_val
60 mean-dummy SD median NA 61 mean-dummy SD median key_val 62 mean-dummy SD mode NA 63 mean-dummy SD mode key_val 64 mean-dummy IQR mean NA 65 mean-dummy IQR mean key_val 66 mean-dummy IQR median NA 67 mean-dummy IQR median NA 68 mean-dummy IQR mode NA 69 mean-dummy IQR mode NA 69 mean-dummy IQR mode key_val 70 median-dummy none NA NA 71 median-dummy none NA key_val 72 median-dummy SD mean NA 73 median-dummy SD mean NA 74 median-dummy SD median NA 75 median-dummy SD median NA 76 median-dummy SD median NA 77 median-dummy SD mode NA 78 median-dummy SD mode NA 79 median-dummy IQR mean NA 80 median-dummy IQR mean NA	58	mean-dummy	SD	mean	NA
61mean-dummySDmediankey_val62mean-dummySDmodeNA63mean-dummySDmodekey_val64mean-dummyIQRmeanNA65mean-dummyIQRmedianNA66mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodeNA78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmeankey_val	59	mean-dummy	SD	mean	key_val
62mean-dummySDmodeNA63mean-dummySDmodekey_val64mean-dummyIQRmeanNA65mean-dummyIQRmeankey_val66mean-dummyIQRmedianNA67mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmeanNA		mean-dummy			NA
63mean-dummySDmodekey_val64mean-dummyIQRmeanNA65mean-dummyIQRmeankey_val66mean-dummyIQRmedianNA67mean-dummyIQRmodeNA68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA75median-dummySDmedianNA75median-dummySDmodeNA76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		mean-dummy	SD	median	key_val
64mean-dummyIQRmeanNA65mean-dummyIQRmeankey_val66mean-dummyIQRmedianNA67mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummysDmeanNA72median-dummySDmeankey_val74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmeanNA		mean-dummy		mode	NA
65 mean-dummy IQR mean key_val 66 mean-dummy IQR median NA 67 mean-dummy IQR median key_val 68 mean-dummy IQR mode NA 69 mean-dummy IQR mode key_val 70 median-dummy none NA NA 71 median-dummy none NA key_val 72 median-dummy SD mean NA 73 median-dummy SD mean key_val 74 median-dummy SD median NA 75 median-dummy SD median NA 76 median-dummy SD median key_val 77 median-dummy SD mode NA 78 median-dummy SD mode NA 79 median-dummy IQR mean NA 79 median-dummy IQR mean NA 80 median-dummy IQR mean NA 80 median-dummy IQR mean NA 80 median-dummy IQR mean NA		mean-dummy		mode	key_val
66mean-dummyIQRmedianNA67mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		mean-dummy	IQR	mean	NA
67mean-dummyIQRmediankey_val68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		mean-dummy			key_val
68mean-dummyIQRmodeNA69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		mean-dummy			NA
69mean-dummyIQRmodekey_val70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA					key_val
70median-dummynoneNANA71median-dummynoneNAkey_val72median-dummySDmeanNA73median-dummySDmedianNA74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		mean-dummy	IQR		NA
71 median-dummy none NA key_val 72 median-dummy SD mean NA 73 median-dummy SD mean key_val 74 median-dummy SD median NA 75 median-dummy SD median key_val 76 median-dummy SD mode NA 77 median-dummy SD mode key_val 78 median-dummy IQR mean NA 79 median-dummy IQR mean NA 80 median-dummy IQR median NA		mean-dummy	IQR		key_val
72median-dummySDmeanNA73median-dummySDmeankey_val74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA			none	NA	NA
73median-dummySDmeankey_val74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		median-dummy		NA	. —
74median-dummySDmedianNA75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA		median-dummy		mean	NA
75median-dummySDmediankey_val76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA	73	median-dummy		mean	key_val
76median-dummySDmodeNA77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA	74	median-dummy	SD	median	NA
77median-dummySDmodekey_val78median-dummyIQRmeanNA79median-dummyIQRmeankey_val80median-dummyIQRmedianNA	75		SD	median	key_val
78 median-dummy IQR mean NA 79 median-dummy IQR mean key_val 80 median-dummy IQR median NA	76			mode	NA
79 median-dummy IQR mean key_val 80 median-dummy IQR median NA				mode	. —
80 median-dummy IQR median NA		<u>~</u>		mean	NA
·		<u>~</u>		mean	
81 median-dummy IQR median key_val		<u>~</u>			
	81	median-dummy	IQR	median	key_val}

82	median-dummy	IQR	mode	NA	
83	median-dummy	IQR	mode	key_val	
84	mode-dummy	none	NA	NA	
85	mode-dummy	none	NA	key_val	
86	mode-dummy	SD	mean	NA	
87	mode-dummy	SD	mean	key_val	
88	mode-dummy	SD	median	NA	
89	mode-dummy	SD	median	key_val	
90	mode-dummy	SD	mode	\overline{NA}	
91	mode-dummy	SD	mode	key_val	
92	mode-dummy	IQR	mean	NA	
93	mode-dummy	IQR	mean	key_val	
94	mode-dummy	IQR	median	NA	
95	mode-dummy	IQR	median	key_val	
96	mode-dummy	IQR	mode	NA	
97	mode-dummy	IQR	mode	key_val	

Table 1: Numbering of cleaning pipelines for Airbnb dataset.