

**Reduced dataset for question 1**

Obs	SITE	YEAR	N	missing	percent_missing
1	15	2003	256	2	0.781
2	19	2005	365	1	0.274
3	20	2005	365	1	0.274
4	23	2006	257	1	0.389
5	31	2003	365	8	2.192
6	31	2005	365	2	0.548
7	32	2003	214	8	3.738
8	32	2005	365	2	0.548
9	35	2004	366	1	0.273
10	35	2005	365	10	2.740
11	36	2003	148	148	100.000
12	36	2004	366	366	100.000
13	36	2005	365	365	100.000
14	36	2006	365	365	100.000
15	43	2007	365	2	0.548
16	45	2009	365	1	0.274
17	45	2010	365	20	5.479
18	46	2007	365	1	0.274
19	46	2008	366	1	0.273
20	46	2009	365	1	0.274

**Final dataset for question 3**

Obs	SITE	YEAR	_TYPE_	_FREQ_	old_mean	new_mean
1	15	2003	0	256	11.984645669	11.984645669
2	19	2005	0	365	7.1101648352	7.1101648352
3	20	2005	0	365	7.1101648352	7.1101648352
4	23	2006	0	257	5.2767578125	5.2767578125
5	31	2003	0	365	7.9907647059	7.9907647059
6	31	2005	0	365	8.2309586777	8.2309586777
7	32	2003	0	214	11.337859223	11.337859223
8	32	2005	0	365	8.2309586777	8.2309586777
9	35	2004	0	366	10.53260274	10.53260274
10	35	2005	0	365	10.605352113	10.605352113
11	36	2003	0	148	.	9.0692567568
12	36	2004	0	366	.	8.5985843373
13	36	2005	0	365	.	8.7591304348
14	36	2006	0	365	.	9.1698055556
15	43	2007	0	365	6.4639118457	6.4639118457
16	45	2009	0	365	10.06043956	10.06043956
17	45	2010	0	365	9.6875362319	9.6871014493
18	46	2007	0	365	10.860164835	10.860164835
19	46	2008	0	366	10.06739726	10.06739726
20	46	2009	0	365	10.06043956	10.06043956