Data Frame Summary

csv_input

Dimensions: 1000 x 78
Duplicates: 0

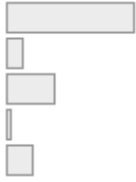| No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
1	X [integer]	Mean (sd) : 499.5 (288.819) min < med < max: 0 < 499.5 < 999 IQR (CV) : 499.5 (0.578)	1000 distinct values (Integer sequence)		1000 (100.0%)	0 (0.0%)
2	case_id [character]	1. 0067a0e0-63d8-11e8-bcf1-0 2. 00dc49de-1fba-11e9-b7f8-0 3. 010df72d-63d9-11e8-bcf1-0 4. 01b2691b-63d8-11e8-bcf1-0 5. 0232701d-6d00-440c-af6c-5 6. 024cad7d-63d9-11e8-bcf1-0 7. 02b84fb5-1fba-11e9-b7f8-0 8. 02f4db2f-63d8-11e8-bcf1-0 9. 038a0679-63d9-11e8-bcf1-0 10. 043a5f9d-63d8-11e8-bcf1-0 [990 others]	1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 990 (99.0%)		1000 (100.0%)	0 (0.0%)
3	case_submitter_id [character]	1. pooled sample 2. QC1 3. QC5 4. QC6 5. WHIM-P93 6. WHIM-P94 7. 01BR001 8. 01BR008 9. 01BR009 10. 01BR010 [984 others]	2 (0.2%) 2 (0.2%) 2 (0.2%) 2 (0.2%) 2 (0.2%) 2 (0.2%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 984 (98.4%)		1000 (100.0%)	0 (0.0%)
4	status [character]	1. Disqualified 2. Qualified	9 (0.9%) 991 (99.1%)		1000 (100.0%)	0 (0.0%)
5	disease_type [character]	1. Breast Invasive Carcinoma 2. Ovarian Serous Cystadenoc 3. Lung Adenocarcinoma 4. Lung Squamous Cell Carcin 5. Glioblastoma 6. Other 7. Colon Adenocarcinoma 8. Pancreatic Ductal Adenoca 9. Uterine Corpus Endometria 10. Clear Cell Renal Cell Car [3 others]	198 (19.8%) 164 (16.4%) 118 (11.8%) 118 (11.8%) 100 (10.0%) 88 (8.8%) 83 (8.3%) 56 (5.6%) 27 (2.7%) 23 (2.3%) 25 (2.5%)		1000 (100.0%)	0 (0.0%)
6	primary_site [character]	1. Brain 2. Breast 3. Bronchus and lung 4. Colon 5. Head and Neck 6. Kidney 7. Lung 8. Ovary 9. Pancreas 10. Uterus, NOS	110 (11.8%) 198 (21.3%) 133 (14.3%) 83 (8.9%) 23 (2.5%) 25 (2.7%) 110 (11.8%) 164 (17.7%) 56 (6.0%) 27 (2.9%)		929 (92.9%)	71 (7.1%)

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
7	demographic_id [character]	1. 00a81a38-1aa2-4498-8778-f 2. 01508809-991b-4fb0-8478-a 3. 0375bfbc-219f-4704-9e14-d 4. 03af35c2-0271-4be6-bdbf-9 5. 05602e4d-4406-44e9-989a-d 6. 094e354d-b9f3-438a-b686-5 7. 0b9f92f2-41db-4500-8964-f 8. 0bf72a4a-f412-479f-82ef-4 9. 0f2ef1b4-65d2-422f-a664-7 10. 0f50346c-89a5-48e7-ac86-5 [990 others]	1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 990 (99.0%)		1000 (100.0%)	0 (0.0%)
8	demographic_submitter_id [character]	1. pooled sample-DM 2. 01BR001-DM 3. 01BR008-DM 4. 01BR009-DM 5. 01BR010-DM 6. 01BR015-DM 7. 01BR017-DM 8. 01BR018-DM 9. 01BR020-DM 10. 01BR023-DM [989 others]	2 (0.2%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 989 (98.9%)		1000 (100.0%)	0 (0.0%)
9	ethnicity [character]	1. Hispanic or Latino 2. Not Hispanic or Latino	22 (4.5%) 468 (95.5%)		490 (49.0%)	510 (51.0%)
10	gender [character]	1. Female 2. Male 3. Unspecified	580 (64.3%) 318 (35.3%) 4 (0.4%)		902 (90.2%)	98 (9.8%)
11	race [character]	1. American Indian or Alaska 2. Asian 3. Black or African American 4. Unknown 5. White	4 (0.7%) 62 (11.5%) 37 (6.9%) 3 (0.6%) 431 (80.3%)		537 (53.7%)	463 (46.3%)
12	vital_status [character]	1. alive 2. Alive 3. Dead	1 (0.2%) 309 (67.0%) 151 (32.8%)		461 (46.1%)	539 (53.9%)
13	days_to_birth [logical]	All NA's			0 (0.0%)	1000 (100.0%)
14	days_to_death [logical]	All NA's			0 (0.0%)	1000 (100.0%)
15	year_of_birth [integer]	Mean (sd) : 1886.097 (353.775) min < med < max: 0 < 1952 < 1993 IQR (CV) : 14 (0.188)	64 distinct values		590 (59.0%)	410 (41.0%)
16	year_of_death [integer]	Mean (sd) : 1343.438 (952.447) min < med < max: 0 < 2016.5 < 2020 IQR (CV) : 2018 (0.709)	15 distinct values		192 (19.2%)	808 (80.8%)
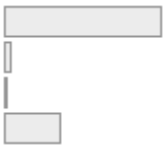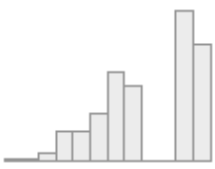
No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
17	diagnosis_id [character]	1. 000e96ce-736f-489f-98a4-3 2. 0555dde7-0be3-478a-8ec2-5 3. 05eed687-3bbd-4ce0-8099-7 4. 0761d368-e6ac-4691-b53d-b 5. 0766e331-52d1-4d4a-8f36-5 6. 0833dc7e-0fbc-4ee5-8385-5 7. 0b1d58a7-a715-4f31-9f69-1 8. 0dbc1dd0-3696-4577-925d-7 9. 0e80a5f2-0f5b-11e9-a064-0 10. 0fa8f8b2-3f24-4a08-8f73-a [990 others]	1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 990 (99.0%)		1000 (100.0%)	0 (0.0%)
18	diagnosis_submitter_id [character]	1. pooled sample-DX 2. 01BR001-DX 3. 01BR008-DX 4. 01BR009-DX 5. 01BR010-DX 6. 01BR015-DX 7. 01BR017-DX 8. 01BR018-DX 9. 01BR020-DX 10. 01BR023-DX [989 others]	2 (0.2%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 1 (0.1%) 989 (98.9%)		1000 (100.0%)	0 (0.0%)
19	age_at_diagnosis [integer]	Mean (sd) : 19210.7 (8893.556) min < med < max: 0 < 21900 < 32872 IQR (CV) : 7636 (0.463)	593 distinct values		809 (80.9%)	191 (19.1%)
20	classification_of_tumor [character]	1. Chromophobe RCC 2. Clear cell RCC 3. Papillary RCC 4. Primary	1 (0.9%) 4 (3.4%) 1 (0.9%) 111 (94.9%)		117 (11.7%)	883 (88.3%)
21	days_to_last_follow_up [integer]	Mean (sd) : 437.256 (642.873) min < med < max: -9 < 264 < 5481 IQR (CV) : 657 (1.47)	348 distinct values		673 (67.3%)	327 (32.7%)
22	days_to_last_known_disease_status [integer]	Mean (sd) : 249.64 (330.558) min < med < max: -1 < 6 < 1425 IQR (CV) : 440 (1.324)	221 distinct values		509 (50.9%)	491 (49.1%)
23	days_to_recurrence [integer]	Mean (sd) : 56.764 (161.29) min < med < max: 0 < 0 < 1034 IQR (CV) : 0 (2.841)	49 distinct values		331 (33.1%)	669 (66.9%)
24	last_known_disease_status [character]	1. Tumor free 2. Tumor Free 3. Unknown tumor status 4. Unknown Tumor Status 5. With tumor 6. With Tumor	52 (16.4%) 98 (30.9%) 11 (3.5%) 40 (12.6%) 34 (10.7%) 82 (25.9%)		317 (31.7%)	683 (68.3%)

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
25	morphology [character]	1. 8500/3 2. 8441/3 3. 8070/3 4. 8140/3 5. 9440/3 6. 8380/3 7. 8312/3 8. 8520/3 9. 8070/33 10. 8140/33 [6 others]	212 (28.5%) 147 (19.8%) 110 (14.8%) 105 (14.1%) 100 (13.5%) 26 (3.5%) 19 (2.6%) 16 (2.2%) 1 (0.1%) 1 (0.1%) 6 (0.8%)		743 (74.3%)	257 (25.7%)
26	primary_diagnosis [character]	1. Infiltrating duct carcino 2. Squamous cell carcinoma, 3. Adenocarcinoma, NOS 4. Glioblastoma 5. Serous adenocarcinoma, NO 6. Adenocarcinoma 7. Serous cystadenocarcinoma 8. Endometrioid adenocarcino 9. Renal cell carcinoma, NOS 10. Infiltrating lobular carc [11 others]	215 (24.8%) 111 (12.8%) 107 (12.3%) 100 (11.5%) 95 (11.0%) 94 (10.8%) 63 (7.3%) 26 (3.0%) 19 (2.2%) 10 (1.2%) 27 (3.1%)		867 (86.7%)	133 (13.3%)
27	progression_or_recurrence [character]	1. no 2. No 3. yes 4. Yes	39 (15.4%) 136 (53.8%) 2 (0.8%) 76 (30.0%)		253 (25.3%)	747 (74.7%)
28	site_of_resection_or_biopsy [character]	1. Lung, NOS 2. Breast, NOS 3. Upper lobe, lung 4. Ovary 5. Lower lobe, lung 6. Head of pancreas 7. Frontal Lobe 8. Temporal Lobe 9. Kidney, NOS 10. Corpus uteri [14 others]	130 (22.3%) 67 (11.5%) 58 (9.9%) 57 (9.8%) 40 (6.9%) 38 (6.5%) 35 (6.0%) 34 (5.8%) 19 (3.3%) 15 (2.6%) 90 (15.4%)		583 (58.3%)	417 (41.7%)
29	tissue_or_organ_of_origin [character]	1. Breast, NOS 2. Lung, NOS 3. Ovary 4. Colon, NOS 5. Brain, NOS 6. Lower lobe, lung 7. Head of pancreas 8. Upper lobe, lung 9. Fallopian tube 10. Frontal lobe [20 others]	191 (20.9%) 160 (17.5%) 131 (14.4%) 83 (9.1%) 44 (4.8%) 44 (4.8%) 38 (4.2%) 30 (3.3%) 22 (2.4%) 22 (2.4%) 147 (16.1%)		912 (91.2%)	88 (8.8%)
30	tumor_grade [character]	1. G1 2. G2 3. G3 4. GX 5. High Grade	13 (5.4%) 137 (57.1%) 82 (34.2%) 7 (2.9%) 1 (0.4%)		240 (24.0%)	760 (76.0%)

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
31	tumor_stage [character]	1. Stage IIA 2. Stage IIB 3. Stage IIIC 4. Stage IIIA 5. IA 6. Stage IB 7. Stage III 8. Stage IIIB 9. IB 10. Stage I [25 others]	136 (19.4%) 89 (12.7%) 75 (10.7%) 59 (8.4%) 48 (6.8%) 39 (5.6%) 37 (5.3%) 36 (5.1%) 35 (5.0%) 33 (4.7%) 115 (16.4%)		702 (70.2%)	298 (29.8%)
32	tumor_cell_content [character]	1. >90 2. 60	5 (83.3%) 1 (16.7%)		6 (0.6%)	994 (99.4%)
33	tumor_largest_dimension_diameter [logical]	All NA's			0 (0.0%)	1000 (100.0%)
34	prior_malignancy [character]	1. No 2. Yes	359 (96.5%) 13 (3.5%)		372 (37.2%)	628 (62.8%)
35	ajcc_clinical_m [character]	1. M0 2. M1 3. Mx 4. Unknown	95 (81.9%) 1 (0.9%) 1 (0.9%) 19 (16.4%)		116 (11.6%)	884 (88.4%)
36	ajcc_clinical_n [character]	1. N0 2. N1 3. N2	87 (84.5%) 6 (5.8%) 10 (9.7%)		103 (10.3%)	897 (89.7%)
37	ajcc_clinical_stage [logical]	All NA's			0 (0.0%)	1000 (100.0%)
38	ajcc_clinical_t [character]	1. pT1 2. pT1a 3. pT3a 4. pT4	1 (16.7%) 1 (16.7%) 3 (50.0%) 1 (16.7%)		6 (0.6%)	994 (99.4%)
39	ajcc_pathologic_m [character]	1. M0 2. M1 3. Unknown	82 (75.9%) 2 (1.9%) 24 (22.2%)		108 (10.8%)	892 (89.2%)
40	ajcc_pathologic_n [character]	1. N0 2. N1 3. N2 4. N0 (i-) 5. N1a 6. N0 (i+) 7. NX 8. N1b 9. N1mi 10. N2a [2 others]	104 (47.9%) 49 (22.6%) 25 (11.5%) 12 (5.5%) 7 (3.2%) 4 (1.8%) 4 (1.8%) 3 (1.4%) 3 (1.4%) 3 (1.4%) 3 (1.4%)		217 (21.7%)	783 (78.3%)

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
41	ajcc_pathologic_stage [character]	1. Stage IIA 2. Stage IIB 3. Stage IIIA 4. Stage IIIB 5. Stage III 6. Stage IB 7. Stage I 8. Stage IA 9. Stage IIIC 10. Stage IV [2 others]	111 (33.1%) 68 (20.3%) 38 (11.3%) 30 (9.0%) 21 (6.3%) 15 (4.5%) 14 (4.2%) 14 (4.2%) 13 (3.9%) 4 (1.2%) 7 (2.1%)		335 (33.5%)	665 (66.5%)
42	ajcc_pathologic_t [character]	1. T1 2. T1a 3. T1b 4. T1c 5. T2 6. T2a 7. T2b 8. T3 9. T3A 10. T4	5 (2.3%) 6 (2.8%) 15 (6.9%) 13 (6.0%) 49 (22.6%) 52 (24.0%) 18 (8.3%) 55 (25.3%) 1 (0.5%) 3 (1.4%)		217 (21.7%)	783 (78.3%)
43	ann_arbor_b_symptoms [logical]	All NA's			0 (0.0%)	1000 (100.0%)
44	ann_arbor_clinical_stage [logical]	All NA's			0 (0.0%)	1000 (100.0%)
45	ann_arbor_extranodal_involvement [logical]	All NA's			0 (0.0%)	1000 (100.0%)
46	ann_arbor_pathologic_stage [logical]	All NA's			0 (0.0%)	1000 (100.0%)
47	best_overall_response [logical]	All NA's			0 (0.0%)	1000 (100.0%)
48	burkitt_lymphoma_clinical_variant [logical]	All NA's			0 (0.0%)	1000 (100.0%)
49	cause_of_death [character]	1. Cancer Related 2. Infection 3. Not Cancer Related 4. Surgical Complications 5. Unknown	24 (57.1%) 3 (7.1%) 9 (21.4%) 1 (2.4%) 5 (11.9%)		42 (4.2%)	958 (95.8%)
50	circumferential_resection_margin [logical]	All NA's			0 (0.0%)	1000 (100.0%)
51	colon_polyps_history [logical]	All NA's			0 (0.0%)	1000 (100.0%)
52	days_to_best_overall_response [logical]	All NA's			0 (0.0%)	1000 (100.0%)
53	days_to_diagnosis [integer]	1 distinct value	0 : 132 (100.0%)		132 (13.2%)	868 (86.8%)
54	days_to_hiv_diagnosis [logical]	All NA's			0 (0.0%)	1000 (100.0%)
55	days_to_new_event [logical]	All NA's			0 (0.0%)	1000 (100.0%)

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
56	figo_stage [character]	1. Stage IC 2. Stage IIA 3. Stage IIC 4. Stage III 5. Stage IIIA 6. Stage IIIB 7. Stage IIIC 8. Stage IV	5 (3.6%) 1 (0.7%) 6 (4.3%) 1 (0.7%) 2 (1.4%) 8 (5.8%) 91 (65.9%) 24 (17.4%)		138 (13.8%)	862 (86.2%)
57	hiv_positive [logical]	All NA's			0 (0.0%)	1000 (100.0%)
58	hpv_positive_type [logical]	All NA's			0 (0.0%)	1000 (100.0%)
59	hpv_status [logical]	All NA's			0 (0.0%)	1000 (100.0%)
60	iss_stage [logical]	All NA's			0 (0.0%)	1000 (100.0%)
61	laterality [logical]	All NA's			0 (0.0%)	1000 (100.0%)
62	ldh_level_at_diagnosis [logical]	All NA's			0 (0.0%)	1000 (100.0%)
63	ldh_normal_range_upper [logical]	All NA's			0 (0.0%)	1000 (100.0%)
64	lymph_nodes_positive [integer]	Mean (sd) : 1.475 (2.602) min < med < max: 0 < 0 < 11 IQR (CV) : 2 (1.765)	0 : 37 (62.7%) 1 : 6 (10.2%) 2 : 2 (3.4%) 3 : 2 (3.4%) 4 : 6 (10.2%) 5 : 1 (1.7%) 7 : 3 (5.1%) 10 : 1 (1.7%) 11 : 1 (1.7%)		59 (5.9%)	941 (94.1%)
65	lymphatic_invasion_present [logical]	All NA's			0 (0.0%)	1000 (100.0%)
66	method_of_diagnosis [logical]	All NA's			0 (0.0%)	1000 (100.0%)
67	new_event_anatomic_site [logical]	All NA's			0 (0.0%)	1000 (100.0%)
68	new_event_type [logical]	All NA's			0 (0.0%)	1000 (100.0%)
69	overall_survival [logical]	All NA's			0 (0.0%)	1000 (100.0%)
70	perineural_invasion_present [logical]	All NA's			0 (0.0%)	1000 (100.0%)
71	prior_treatment [character]	1. No	132 (100.0%)		132 (13.2%)	868 (86.8%)
72	progression_free_survival [logical]	All NA's			0 (0.0%)	1000 (100.0%)
73	progression_free_survival_event [logical]	All NA's			0 (0.0%)	1000 (100.0%)

No	Variable	Stats / Values	Freqs (% of Valid)	Graph	Valid	Missing
74	residual_disease [character]	1. R0 2. R1 3. R2 4. RX	117 (70.5%) 5 (3.0%) 2 (1.2%) 42 (25.3%)		166 (16.6%)	834 (83.4%)
75	vascular_invasion_present [logical]	All NA's			0 (0.0%)	1000 (100.0%)
76	year_of_diagnosis [integer]	Mean (sd) : 2011.163 (5.566) min < med < max: 1994 < 2010 < 2017 IQR (CV) : 9 (0.003)	17 distinct values		258 (25.8%)	742 (74.2%)
77	icd_10_code [logical]	All NA's			0 (0.0%)	1000 (100.0%)
78	synchronous_malignancy [logical]	All NA's			0 (0.0%)	1000 (100.0%)

Generated by [summarytools](#) 0.9.9 ([R](#) version 4.0.3)
2021-05-13