**Mappings of the medical document with the csv file**

1. ID
2. age
3. sex: {male: 1, female: 2}
4. subsite: {tongue: 1, BM: 2, FOM: 3, alveolus: 4, hard palate: 5, RMT: 6, lips: 7}
5. pathology: {WDSCC: 1, Verrucous: 2, MDSCC: 3, PDSCC: 4, others: 5, ca in situ: 6}
6. clinicalT8thed: {T1: 1, T2: 2, T3: 3, T4a: 4, T4b: 5}
7. nodal8thed: {No nodes: 0, N1: 1, N2a: 2, N2b: 3, N2c: 4, N3a: 5, N3b: 6}
8. pt8thed: {T1: 1, T2: 2, T3: 3, T4a: 4, T4b: 5}
9. pN8thed: {N0: 0, N1: 1, N2a: 2, N2b: 3, N2c: 4, N3a: 5, N3b: 6}
10. PNI: {no: 0, yes: 1}
11. LVI: {no: 0, yes: 1}
12. Margins: {negative: 0, close: 1, positive: 2}
13. Boneinv: {no: 0, yes: 1}
14. skininv: {no: 0, yes: 1}
15. ENS: {no: 0, yes: 1}
16. Surgery: {SND 1-3: 1, SND 1-4: 2, MRND: 3, RND: 4}
17. lateralitysurgery: {ipsilateral: 1, bilateral: 2}
18. DOS
19. adjRx: {no: 0, yes: 1}
20. date of LR
21. Nodalrecurrencelevel: {ipsilateral: 1, contralateral: 2}
22. Nodallevel
23. date of Nodalrecurrence
24. last FU
25. stateFU: {NED: 1, AWD: 2, DWD': 3, Dead due to other causes: 4}

**Fullforms**

1. BM – Buccal Mucosa
2. FOM – Floor of Mouth
3. RMT – Retromolar Trigone
4. WDSCC – Well-Differentiated Squamous Cell Carcinoma
5. MDSCC – Moderately Differentiated Squamous Cell Carcinoma
6. PDSCC – Poorly Differentiated Squamous Cell Carcinoma
7. PNI – Perineural Invasion
8. LVI – Lymphovascular Invasion
9. ENS – Extracapsular Nodal Spread
10. SND – Selective Neck Dissection
11. MRND – Modified Radical Neck Dissection
12. RND – Radical Neck Dissection
13. DOS – Date of Surgery
14. adjRx – Adjuvant Therapy
15. date of LR – Date of Local Recurrence
16. last FU – Date of Last Follow-Up
17. state FU – State at Last Follow-Up
18. Verrucous - Verrucous Carcinoma