Counting regions for tissuecyte

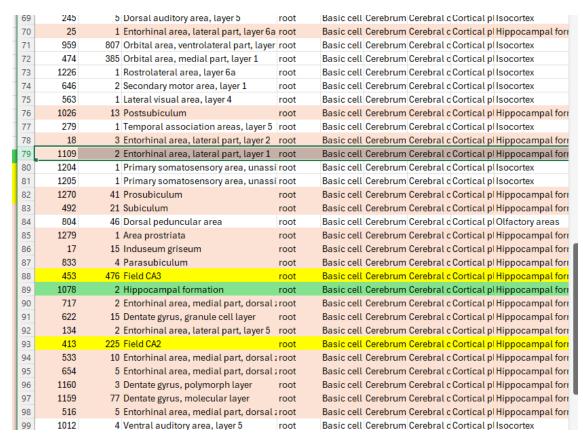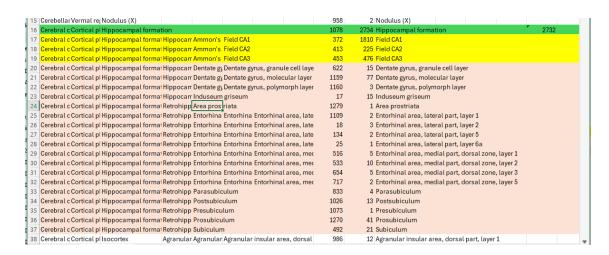| Status	Done
◎ Project	◎ 1_Recent To Do
↗ Parent-task	🗋 U01 -Gopi group
🏷 Tags	

Description

▼ newer discussion on 240827, use B0039, 240826_Gui_registration_second_test result as an example.

- we need to figure out how counting code output files can be converted to 300 summary regions output

- First question:

  - what does region number refer to for cell_counts.csv file? Is it referring to parcellation_index or id in the **1_adult_mouse_brain_graph_mapping_all_reigons file**?

  - for example for region = 1080 in cell_counts.csv file, I found 1080 exist as both parcellation_index and id which refers to different brain regions.

	A	B	C
1	region	count	
2	926	1	
3	1053	1	
4	641	1	
5	1080	4	
6	958	2	

- - I have figured out based on CA1. I believe region number refers to parcellation_index. And the current **cell_region_count.csv** is correct.

- So the **cell_count.csv =** this file store all detected signals and their location (parcellation_index).

- Then based on **1_adult_mouse_brain_graph_mapping_all_reigons file**, it can know the all detected signal which brain regions they belong to. This create the **cell-region_count.csv.** But this file only have counts for all finest regions. It does not sum fine regions to higher level. For example: all yellow and orange are the finest regions (eg: CA1,CA2,CA3, parasubiculum etc.), but green is the higher level region (hippocampal formation = 2). (This means annotation file has hippocampal parcellation index and 2 cells falls in this index.)
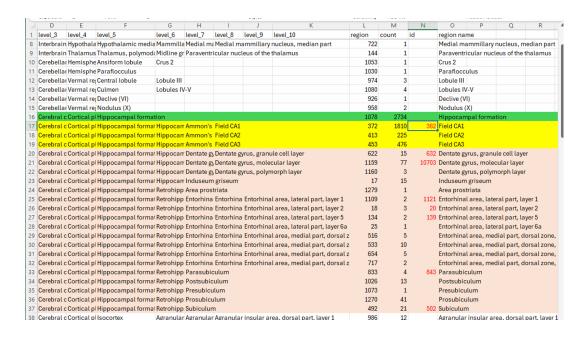
69	245	5	Dorsal auditory area, layer 5	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
70	25	1	Entorhinal area, lateral part, layer 6a	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
71	959	807	Orbital area, ventrolateral part, layer	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
72	474	385	Orbital area, medial part, layer 1	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
73	1226	1	Rostrolateral area, layer 6a	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
74	646	2	Secondary motor area, layer 1	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
75	563	1	Lateral visual area, layer 4	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
76	1026	13	Postsubiculum	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
77	279	1	Temporal association areas, layer 5	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
78	18	3	Entorhinal area, lateral part, layer 2	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
79	1109	2	Entorhinal area, lateral part, layer 1	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
80	1204	1	Primary somatosensory area, unassi	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
81	1205	1	Primary somatosensory area, unassi	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex
82	1270	41	Prosubiculum	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
83	492	21	Subiculum	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
84	804	46	Dorsal peduncular area	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Olfactory areas
85	1279	1	Area prostriata	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
86	17	15	Induseum griseum	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
87	833	4	Parasubiculum	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
88	453	476	Field CA3	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
89	1078	2	Hippocampal formation	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
90	717	2	Entorhinal area, medial part, dorsal ;	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
91	622	15	Dentate gyrus, granule cell layer	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
92	134	2	Entorhinal area, lateral part, layer 5	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
93	413	225	Field CA2	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
94	533	10	Entorhinal area, medial part, dorsal ;	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
95	654	5	Entorhinal area, medial part, dorsal ;	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
96	1160	3	Dentate gyrus, polymorph layer	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
97	1159	77	Dentate gyrus, molecular layer	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
98	516	5	Entorhinal area, medial part, dorsal ;	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Hippocampal forr
99	1012	4	Ventral auditory area, layer 5	root	Basic cell	Cerebrum	Cerebral c	Cortical pl	Isocortex

- **region_count.csv** is the spatially organized version of **cell_region_count.csv.** At the same time, it added lower level regions counts to higher order. For eg: Green = yellow + orange +2 (the 2 cell detected just in hippocampal formation) =2732+2 =2734.
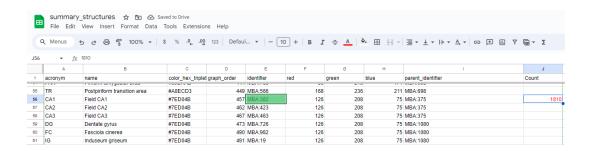
15	Cerebellar Vermal re	Nodulus (X)			958	2	Nodulus (X)	
16	Cerebral c	Cortical pl	Hippocampal formation		1078	2734	Hippocampal formation	2732
17	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Ammon's	Field CA1	372	1810	Field CA1
18	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Ammon's	Field CA2	413	225	Field CA2
19	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Ammon's	Field CA3	453	476	Field CA3
20	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Dentate g	Dentate gyrus, granule cell laye	622	15	Dentate gyrus, granule cell layer
21	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Dentate g	Dentate gyrus, molecular layer	1159	77	Dentate gyrus, molecular layer
22	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Dentate g	Dentate gyrus, polymorph layer	1160	3	Dentate gyrus, polymorph layer
23	Cerebral c	Cortical pl	Hippocampal forma	Hippocan Induseum griseum	17	15	Induseum griseum	
24	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp	Area prostriata	1279	1	Area prostriata
25	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, late	1109	2	Entorhinal area, lateral part, layer 1
26	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, late	18	3	Entorhinal area, lateral part, layer 2
27	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, late	134	2	Entorhinal area, lateral part, layer 5
28	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, late	25	1	Entorhinal area, lateral part, layer 6a
29	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, med	516	5	Entorhinal area, medial part, dorsal zone, layer 1
30	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, med	533	10	Entorhinal area, medial part, dorsal zone, layer 2
31	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, med	654	5	Entorhinal area, medial part, dorsal zone, layer 3
32	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Entorhina Entorhina	Entorhinal area, med	717	2	Entorhinal area, medial part, dorsal zone, layer 5
33	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Parasubiculum	833	4	Parasubiculum	
34	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Postsubiculum	1026	13	Postsubiculum	
35	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Presubiculum	1073	1	Presubiculum	
36	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Prosubiculum	1270	41	Prosubiculum	
37	Cerebral c	Cortical pl	Hippocampal forma	Retrohipp Subiculum	492	21	Subiculum	
38	Cerebral c	Cortical pl	Isocortex	Agranular Agranular Agranular insular area, dorsal	986	12	Agranular insular area, dorsal part, layer 1	

- Now we want convert **region_count.csv** to 300 summary regions file.

- for all the regions in **region_count.csv,** based on their region number/parcellation_index, find their id number from **1_adult_mouse_brain_graph_mapping_all_reigons file.**

- eg: red numbers are the ids that i found manually

- for some regions, like CA1, id =382, it can also be found in 300 **summary_structures.csv.** Put the count number (1810) for id (382) in 300 summary_structures.csv.
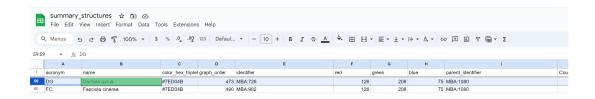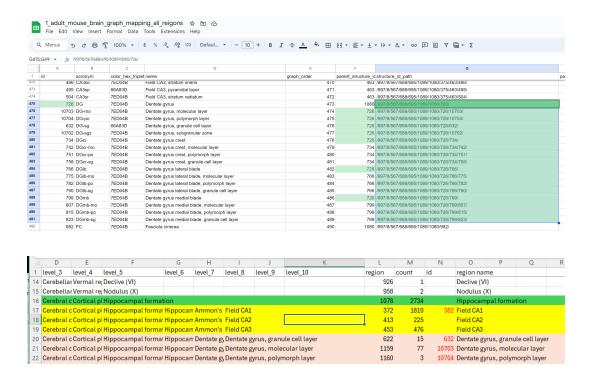

- for some other regions, like dentate gyrus, granule cell layer, region number/parcellation_index =622, id=632. It can not be found in 300 **summary_structures.csv,** because this region level is smaller than the 300 regions we want**.** We need to figure out for this region **(**dentate gyrus, granule cell layer, region number/parcellation_index =622, id=632) is the lower level region of which regions inside **300**

summary_structures.csv. Or in another way, we need to know for the 300 summary_structures, what are the subregions they have. Once we know that info, we can know id=632 belongs to which 300 summary regions.

- we need to first go over each row inside 300 **summary regions.csv,** based on the id number, check the column **structure_id_path** in **1_adult_mouse_brain_graph_mapping_all_reigons.csv.**

- as long as the **structure_id_path** has the id number, it means those rows belong to that id.

- for eg, see pics below: row 59 in summary_structures.csv, id =726. The column **structure_id_path** in **1_adult_mouse_brain_graph_mapping_all_reigons.csv** that has 726 contains row 475-491. All ids (row 475-491.) covered by these rows belongs to the DG/dentate gyrus/id=726. Therefore, all the counts for these ids (row 475-491) should be counted towards DG/dentate gyrus/id=726.

- if we use B0039 as an example: the 300 summary_structures.csv output file for B0039, DG should have counts = (632/**15** + 10703/**77** +10704/**3**) = **95** cells. [the numbers has a format of id/counts]
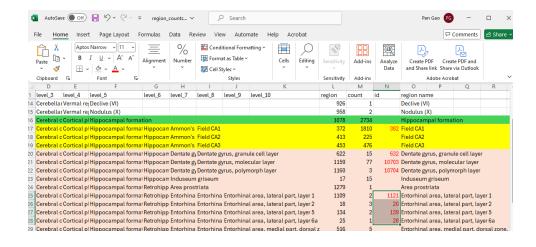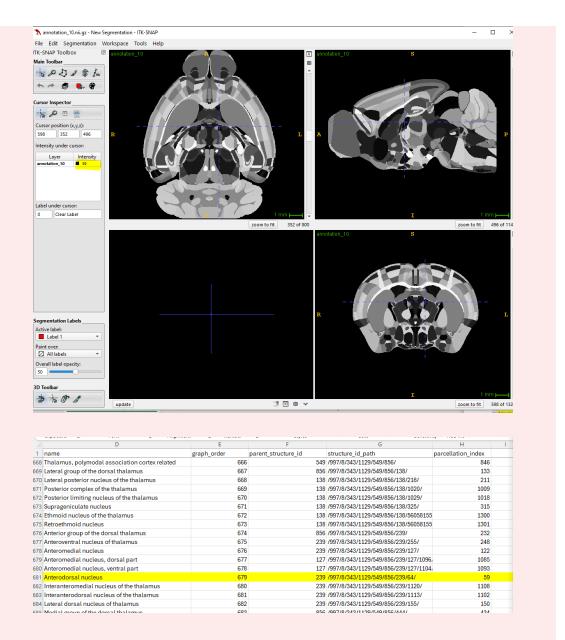
- More example:

if we use B0039 as an example: the 300 summary_structures.csv output file for B0039, ENTI = Entorhinal area = id 918, lateral part should have counts = (1121/2 + 20/3 +139/2 + 28/1) id/counts format = 8 cells.

▼ The steps to find counting number for summarized regions

file location G:\Brain_Stitch\Brain regions and summary regions

1. First create a new column at the end of **all regions file and summary structure files** and call it **counting number** column.

2. Use **all regions file** to find the detected cell number for all regions by using parcellation index number. Put the detected cell number to the **counting number column** for **All regions file**. Do that for all regions in side All regions file.

	D	E	F	G	H	I
1	name	graph_order	parent_structure_id	structure_id_path	parcellation_index	
668	Thalamus, polymodal association cortex related	666		549 /997/8/343/1129/549/856/	846	
669	Lateral group of the dorsal thalamus	667		856 /997/8/343/1129/549/856/138/	133	
670	Lateral posterior nucleus of the thalamus	668		138 /997/8/343/1129/549/856/138/218/	211	
671	Posterior complex of the thalamus	669		138 /997/8/343/1129/549/856/138/1020/	1009	
672	Posterior limiting nucleus of the thalamus	670		138 /997/8/343/1129/549/856/138/1029/	1018	
673	Suprageniculate nucleus	671		138 /997/8/343/1129/549/856/138/325/	315	
674	Ethmoid nucleus of the thalamus	672		138 /997/8/343/1129/549/856/138/56058155	1300	
675	Retroethmoid nucleus	673		138 /997/8/343/1129/549/856/138/56058155	1301	
676	Anterior group of the dorsal thalamus	674		856 /997/8/343/1129/549/856/239/	232	
677	Anteroventral nucleus of thalamus	675		239 /997/8/343/1129/549/856/239/255/	248	
678	Anteromedial nucleus	676		239 /997/8/343/1129/549/856/239/127/	122	
679	Anteromedial nucleus, dorsal part	677		127 /997/8/343/1129/549/856/239/127/1096/	1085	
680	Anteromedial nucleus, ventral part	678		127 /997/8/343/1129/549/856/239/127/1104/	1093	
681	Anterodorsal nucleus	679		239 /997/8/343/1129/549/856/239/64/	59	
682	Interanteromedial nucleus of the thalamus	680		239 /997/8/343/1129/549/856/239/1120/	1108	
683	Interanterodorsal nucleus of the thalamus	681		239 /997/8/343/1129/549/856/239/1113/	1102	
684	Lateral dorsal nucleus of thalamus	682		239 /997/8/343/1129/549/856/239/155/	150	
685	Medial group of the dorsal thalamus	683		856 /997/8/343/1129/549/856/444/	434	

- use the parcellation_index and the id and parent_structure_id and structure_id_path to place correct cell numbers for all regions/all rows.

3. Then look at **summary structure files**, go to first row and identify the **identifier number/ id** (MBA:xxx) for that row. Then use the id to find the correct rows in **All regions file** by checking the **id** column.

(For example for row 119 in **summary structure files**, Anterovenral nucleus of thalamus, the identifier number =255. Go to **All regions file** and the **id** column with 255, which will locate you to row 677. Take the counting

number for that row and place and paste it in **summary structure files, counting number column**)

4. Repeat this step for all rows in **summary structure files.**

5. Output the both the **summary structure files and all regions file.**

1_adult_mouse_brain_graph_mapping.csv

summary_structures.csv