Part 1: Top 10 enriched functions from first 200 genes

Function	Description	Relation to OV
1. Positive regulation of molecular function	Increases molecular functions, including catalysis and binding	Increased cell activity, including signaling, cell division, etc, would be expected in cancer[2]
2. Regulation of hydrolase activity	Controls the frequency or rate of hydrolysis	Hydrolysis is used to break down many types of molecules, for example, lipid metabolism is increased in ovarian cancer[4]
3. Neurogenesis	Generates nerve cells	Tumors can increase neurogenesis, and the nerves have a role in promoting tumor growth and metastasis[5]
4. Neuron projection	Affects transmission of signals in nerve cells	
5. Cytoskeletal protein binding	Interactions with proteins in the cytoskeleton	The cytoskeleton is important in cell division
6. Golgi apparatus	Controls function of golgi apparatus, which is responsible for secretions of ligands and hormones	Deregulated hormone activity and cell-cell signaling is common in OV[1]
7. Regulation of cell differentiation	Controls cell differentiation	Stem cells can be responsible for tumor growth and metastasis[6]
8. Whole membrane	Formation of cell membranes and organelle membranes	Formation of more cell membranes can mean increased cell division
9. Regulation of intracellular signal transduction	Controls signaling between cells	Deregulated cell-cell signaling is a key factor in developing cancer[1]
10. Protein homodimerization activity	Formation of protein homodimers	Protein homodimers are important in signal transduction, and errors in the signal transduction pathways can lead to cancer[3]

1. Chang, C. C., Su, K. M., Lu, K. H., Lin, C. K., Wang, P. H., Li, H. Y., Wang, M. L., Lin, C. K., Yu, M. H., & Chang, C. M. (2018). Key Immunological Functions Involved in the Progression of Epithelial Ovarian Serous Carcinoma Discovered by the Gene Ontology-Based Immunofunctionome Analysis. *International journal of molecular sciences*, *19*(11), 3311. https://doi.org/10.3390/ijms19113311

2. Zhou, Y., Layton, O., & Hong, L. (2018). Identification of Genes and Pathways Involved in Ovarian Epithelial Cancer by Bioinformatics Analysis. *Journal of Cancer*, *9*(17), 3016–3022. https://doi.org/10.7150/jca.26133

3. Sever, R., & Brugge, J. S. (2015). Signal transduction in cancer. *Cold Spring Harbor perspectives in medicine*, *5*(4), a006098. https://doi.org/10.1101/cshperspect.a006098

4. Ji, Z., Shen, Y., Feng, X., Kong, Y., Shao, Y., & Meng, J. et al. (2020). Deregulation of Lipid Metabolism: The Critical Factors in Ovarian Cancer. *Frontiers In Oncology*, *10*. doi: 10.3389/fonc.2020.593017

5. Cervantes-Villagrana, R. D., Albores-García, D., Cervantes-Villagrana, A. R., & García-Acevez, S. J. (2020). Tumor-induced neurogenesis and immune evasion as targets of innovative anti-cancer therapies. *Signal transduction and targeted therapy*, *5*(1), 99. https://doi.org/10.1038/s41392-020-0205-z

6. Kenda Suster, N., & Virant-Klun, I. (2019). Presence and role of stem cells in ovarian cancer. *World journal of stem cells*, *11*(7), 383–397. https://doi.org/10.4252/wjsc.v11.i7.383

Part 2

1)
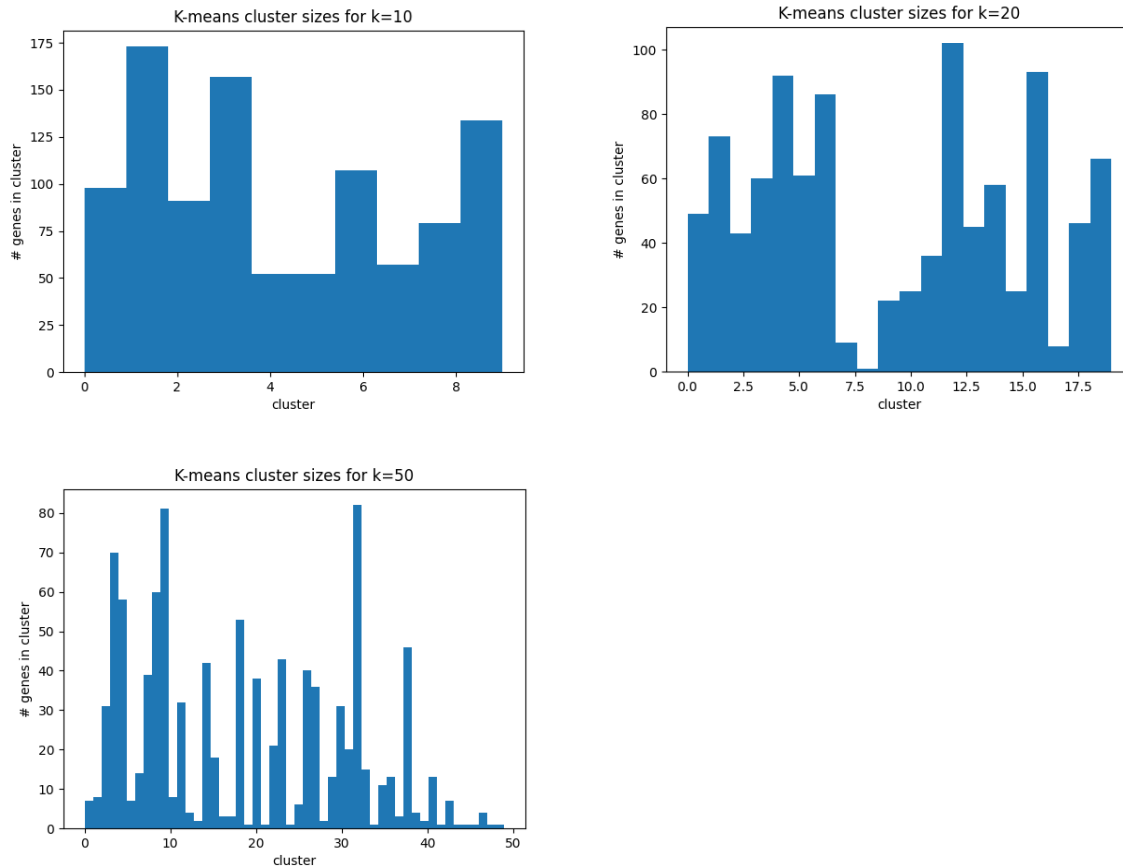
2) Top 10 enriched functions from cluster #2 (k=20): overall, these genes are related to embryo development, including cell differentiation, organ formation, and development of the nervous system.

Genes: CAMK2A, SCRG1, KCNA1, TPO, GLRA2, IL1RAPL2, LBP, PI16, LOC257358, LHFPL5, LRIT3, ST8SIA6, EBF2, TMEM211, KLF17, PDE1C, TREML1, C14orf180, TUBAL3, LOC728606, SLC1A6, ART4, TNFSF18, FADS6, AQPEP, FAM129C, TCL1A, LOC100192378, CTAG1B, TBX5, IGFN1, EPX, CIDEA, CCDC144C, HABP2, TUBB4Q, OR7E91P, ERVFRDE1, FHL5, SLURP1, FOXG1, ISL1, STAB2, SLC1A7, DHH, HAND1, RERGL, GDNF, C3orf36, CCDC60, NOX1, IHH, KCNK4, GPR144, NLRP11, CDHR2, C1orf150, CELP, C14orf184, CCDC144NL, CSMD1, LOXHD1, NPAS4, EFCAB5, SH2D7, ADIPOQ, IFNE, ASPA, TUSC5, HTR2A, SHH, GAGE12D, FBXL21

Top 10 functions:
1. Detection of mechanical stimulus involved in sensory perception: mechanical stimulus converted to molecular signal in nerves
2. Patched binding: receptor for hedgehog protein which is important in embryo development
3. Cell cell signaling: signaling between cells
4. Regulation of animal organ formation: controls organ formation

5. Cell fate commitment: cell differentiation
6. Cell fate specification: same as #5
7. Detection of mechanical stimulus: same as #1
8. Embryo development: formation and development of embryo
9. Animal organ formation: same as #4
10. Neuron fate commitment: cell differentiating to become a neuron

Part 3

1, 2) Accuracy improves using only the top 1000 genes because they are more significant indicators of survival (lower p-values) and using more genes that are less significant introduces more variability, which would decrease the accuracy.

```
All genes:
k = [1, 3, 5]
accuracy = [0.6216216216216216, 0.6486486486486487, 0.6216216216216216]

First 1000 genes:
k = [1, 3, 5]
accuracy = [0.6756756756756757, 0.7027027027027027, 0.7027027027027027]
```

3) Accuracy = 62.16%