Step 1: Glance through [Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors | Nature Genetics](). This is the first published scRNA-seq data of colorectal cancer (CRC). The study sequenced multiple CRC tumours (malignant cells and stromal cells) and normal tissue samples.

Step 2: You can download normalised scRNA-seq data from [GEO Accession viewer (nih.gov)]()

GSE81861_CRC_NM_all_cells_COUNT.csv.gz (normal tissue)

GSE81861_CRC_tumor_all_cells_COUNT.csv.gz (cancer tissue)

Step 3: Use package such as Seurat (you can use something else as well) for data pre-processing, normalisation, clustering etc. Annotate the cells into following categories using a suitable method (Azimut or AUCell). Cells that are unannotated, can be put under "Others" category.

Cell type	Cell-type specific markers
Astrocyte	AGXT2L1, GFAP, ALDOC, SLC1A3, AGT, ALDH1L1
B cell	CD19, MS4A1, BANK1, BLK, IRF8, ABCB4, ABCB9, AFF4, AIDA, AIM2
Endothelial cell	VWF, PECAM1, CDH5, VEGFA, FLT1, ECSCR, ACYP1, ADGRL2, SELE, ICAM1
Epithelial cell	CDH1, MYLK, ANKRD30A, ABCA13, ABCB10, ADGB, SFTPB, SFTPC
Erythrocyte	ALAS2, CA1, HBB, HBE1, HBA1, HBG1, GYPA
Fibroblast	COL1A1, COL3A1, THY1, NECTIN1, FAP, PTPN13, C5AR2, LRP1
GMP	CD38, KIT, ADK, CD123, ALDH4A1, ANXA1, AP3S1, APLP2, APPL1, AREG, ASPM, CDKN3, CLSPN, DEPDC7, MCM10, MUCB2, SDC4, RMI2
HSC	CD34, ITGA5, PROM1, CD105, VCAM1, CD164, THY1, KIT, ACE, CMAH, ABCG2, CD41, ALDH1A1, BMI1
Macro/Mono/DC	CD68, CD14, MRC1, BHLHE40, CD93, CREM, CSF1R, CCL18, ICAM4, ACPP, ACSL3, ADGRE2, ADGRE3, CD209, CD83, CD1A
Malignant cell	EPCAM, FOLH1, KLK3, KRT8, KRT18, KRT19
Mast cell	SLC18A2, ADIRF, ASIC4, BACE2, ENPP3, CADPS, CAPN3, CDK15, CMA1, GCSAML, MAML1, MAOB, CAVIN2
Myeloid cell	PTPRC, CD14, AIF1, TYROBP, CD163
Neuron	STMN2, RBFOX3, MAP2, TUBB3, CSF3, DLG4, ENO2
Neutrophil	ADGRG3, CXCL8, FCGR3B, MNDA, USP10, CSF3R, ANXA3, AQP9, BTNL8, LGALS13, G0S2, NFE4, IL5RA
NK cell	FCGR3A, KLRB1, KLRD1, NKG7, XCL1, XCL2, NCR3, NCR1, CD247, GZMB, KLRC1, KLRK1
Oligodendrocyte	MOG, OLIG1, OLIG2, PDGFRA, PLP1, MBP, MAG, SOX10
Plasma cell	MZB1, BRSK1, AC026202.3, JSRP1, LINC00582, PARM1, TAS1R3
Progenitor	CD38, CASR, ALDH, CAR, KDR, MME, FLT3, CD90, CD123
T cell	CD3D, CD3G, CD3E

(Taken from https://ngdc.cncb.ac.cn/cancerscem/documents)

Step 4: Identify clusters of stromal cells across both normal and cancer datasets.

Step 5: Perform DE analysis for stromal cell types across both normal and cancer datasets. Comment on the extent of phenotypic changes in stromal subtypes due to tumour infiltration. You can do it by pathway analysis or literature survey of cherry picked dysregulated genes.

Step 6: Report which immune cell types are prevalent in cancers. Also report if you see a variability in frequency of cancer infiltrated immune cell types.