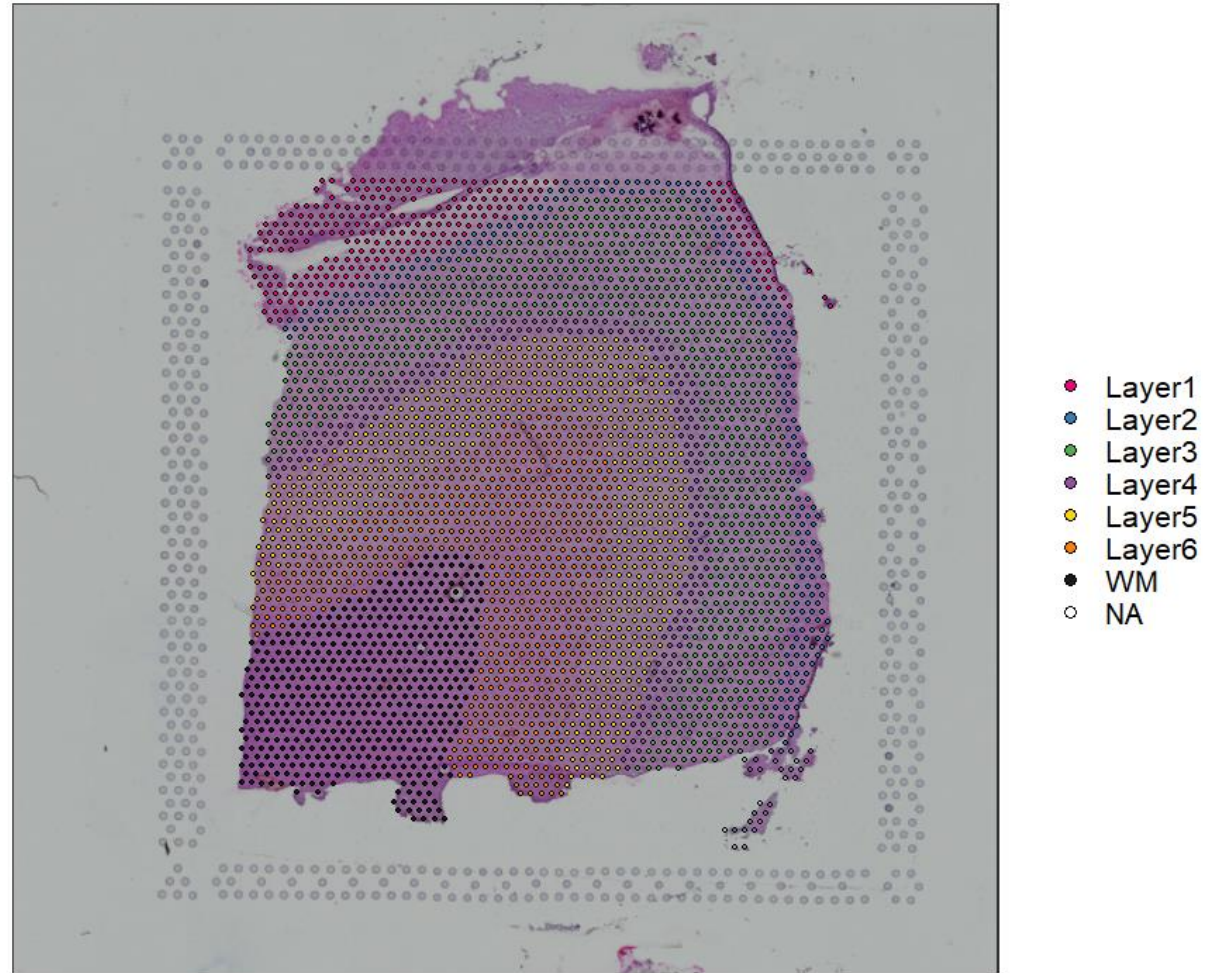


# Analysis of Maynard 2020 brain spatial transcriptomic data

4/1

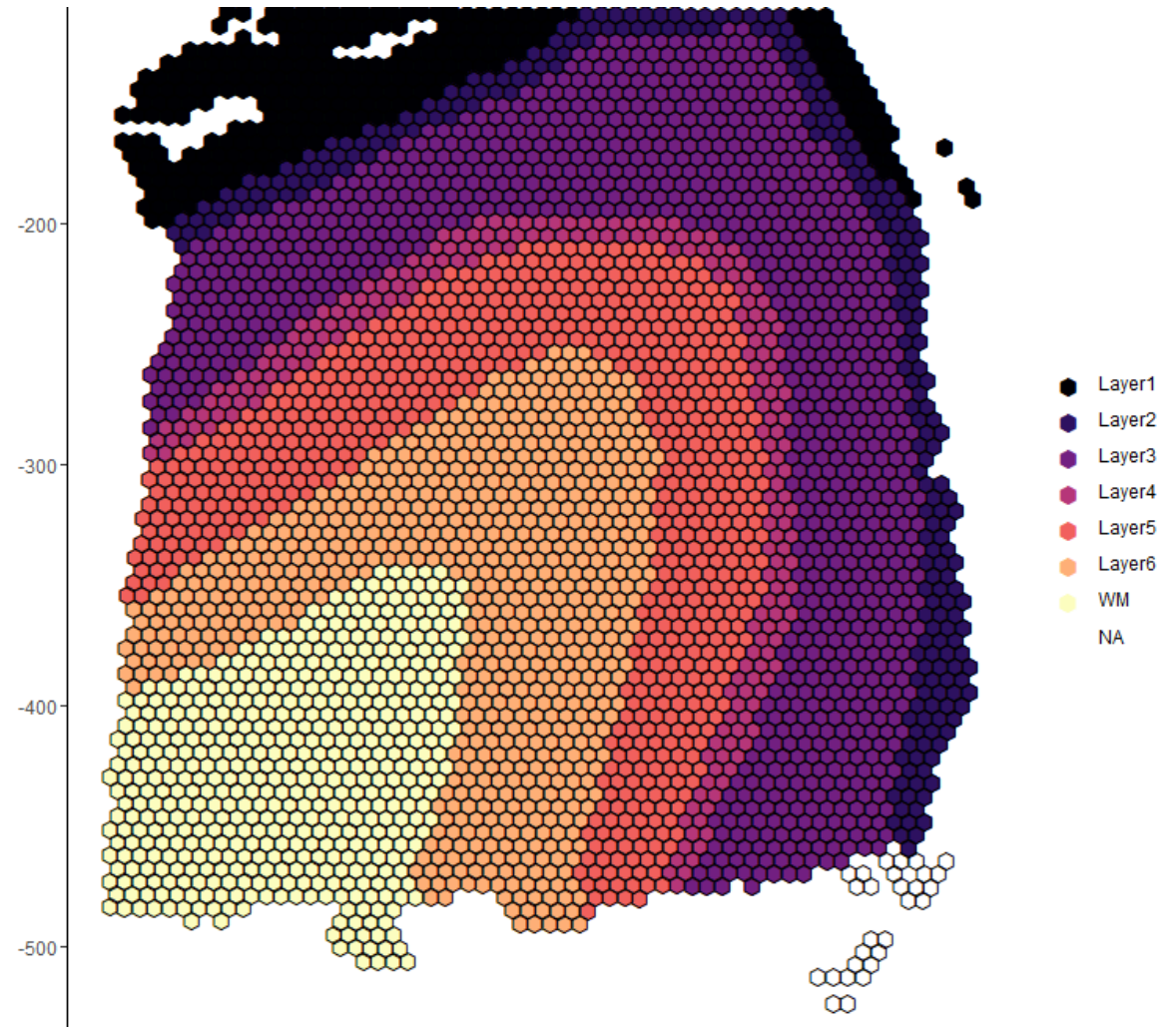
# Sample 151673

- Dorsolateral prefrontal cortex layers
- 3639 capture locations



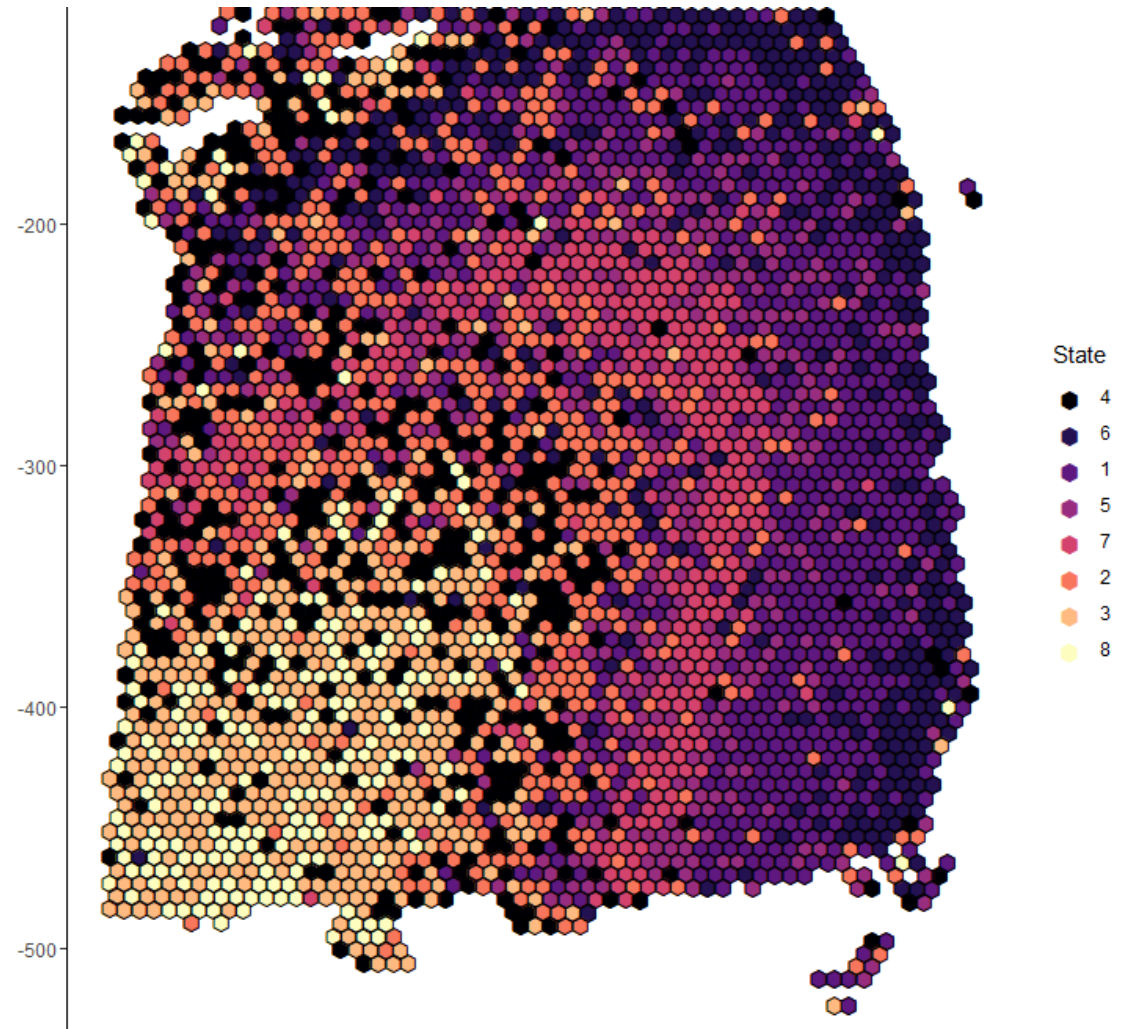
# Ground truth

- Supervised annotation based on “cytoarchitecture and selected gene markers”



# Marker-guided clustering

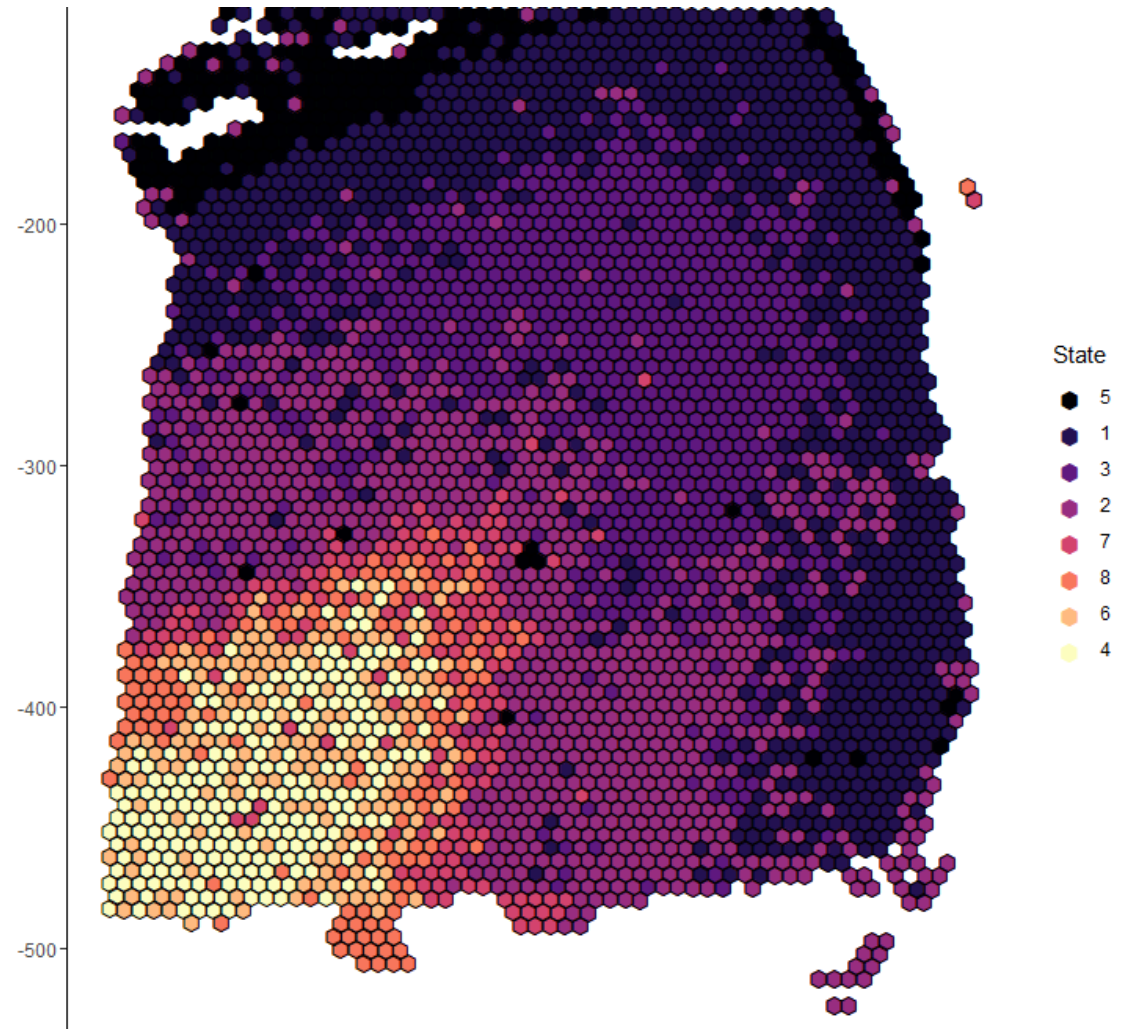
- Unsupervised clustering on PCs from known rodent and human layer marker genes from Zeng et al., 2012 + spatial coordinates
- 8 clusters since it “gave slightly improved clustering performance”
- Adjusted Rand Index (ARI) = 0.197





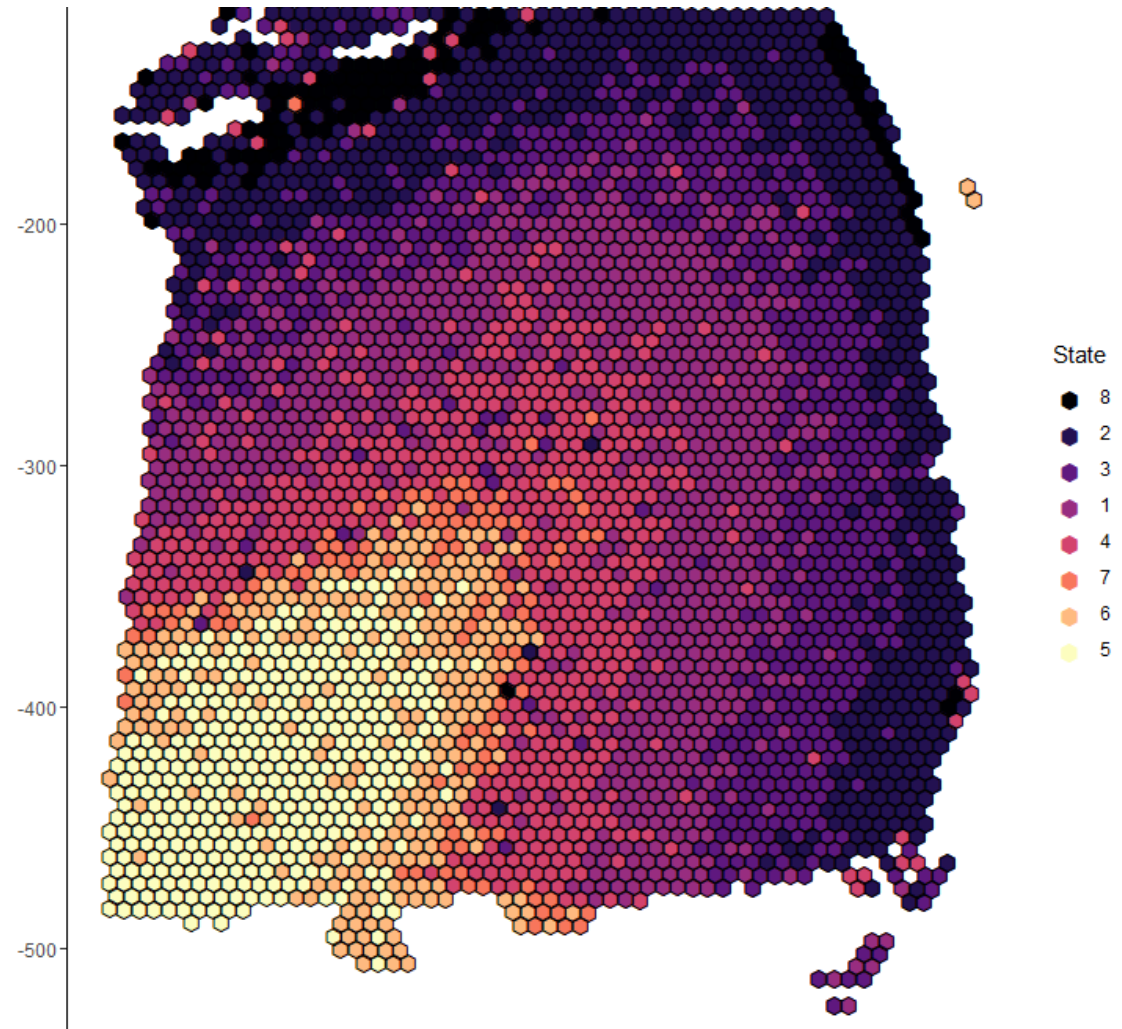
# Unsupervised clustering

- 50 PCs from highly variable genes + spatial coordinates
- 8 clusters
- ARI = 0.274



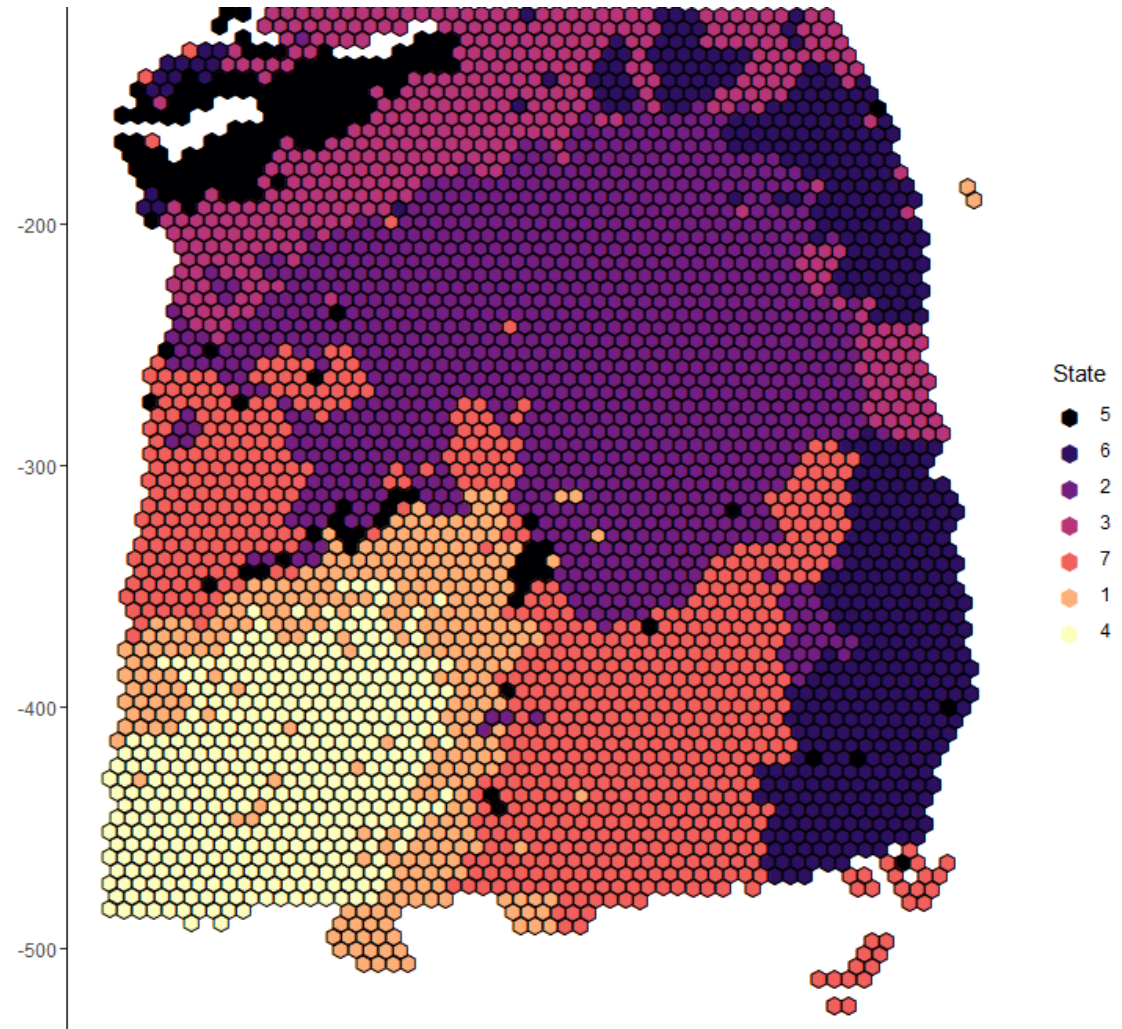
# Semi-supervised clustering

- Unsupervised clustering on PCs from genes differentially expressed by layer + spatial coordinates
  - Circular
- 8 clusters
- ARI = 0.380



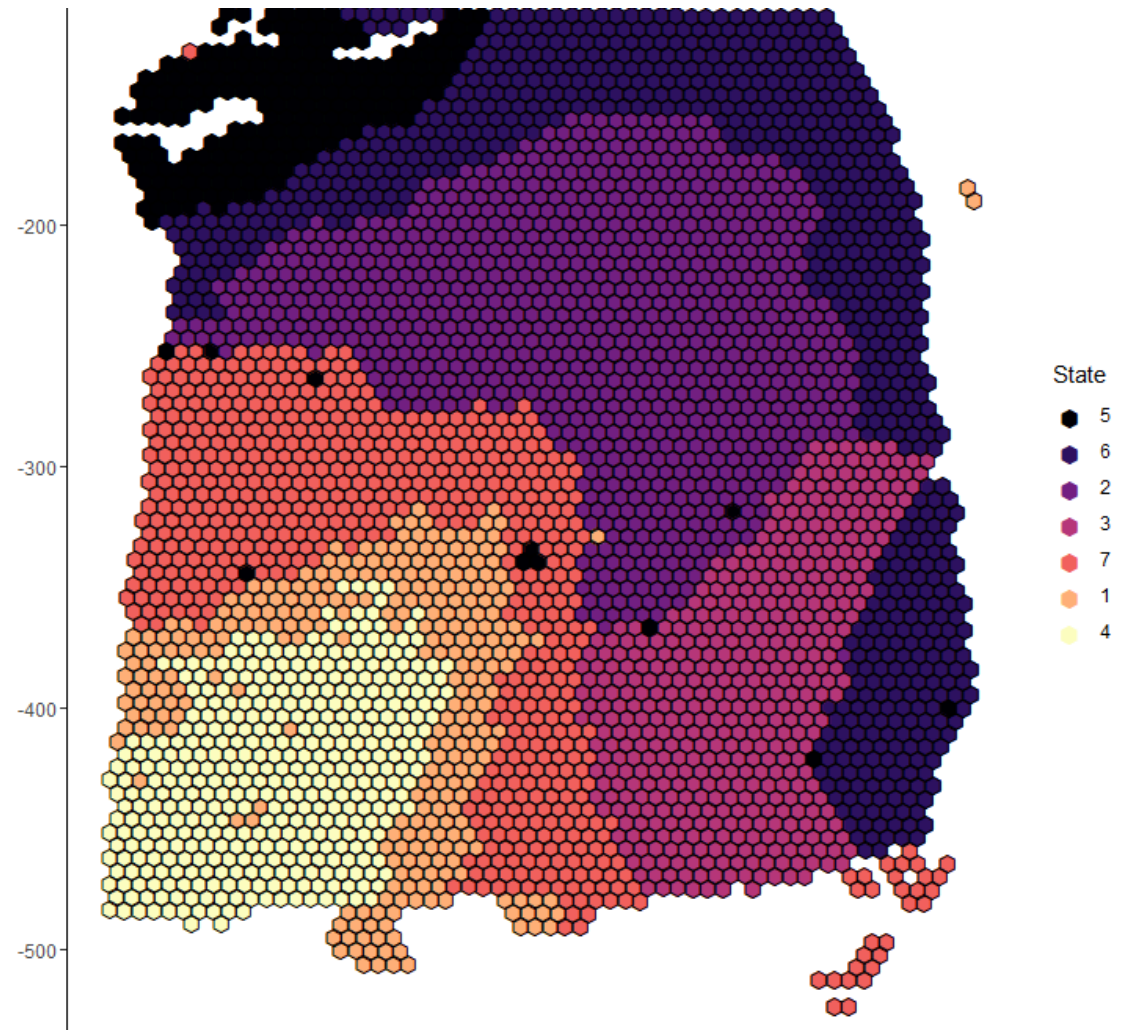
# MCMC Clustering

- 7 clusters
- 5 PCs
- Smoothing parameter = 2
- ARI = 0.242



# MCMC Clustering

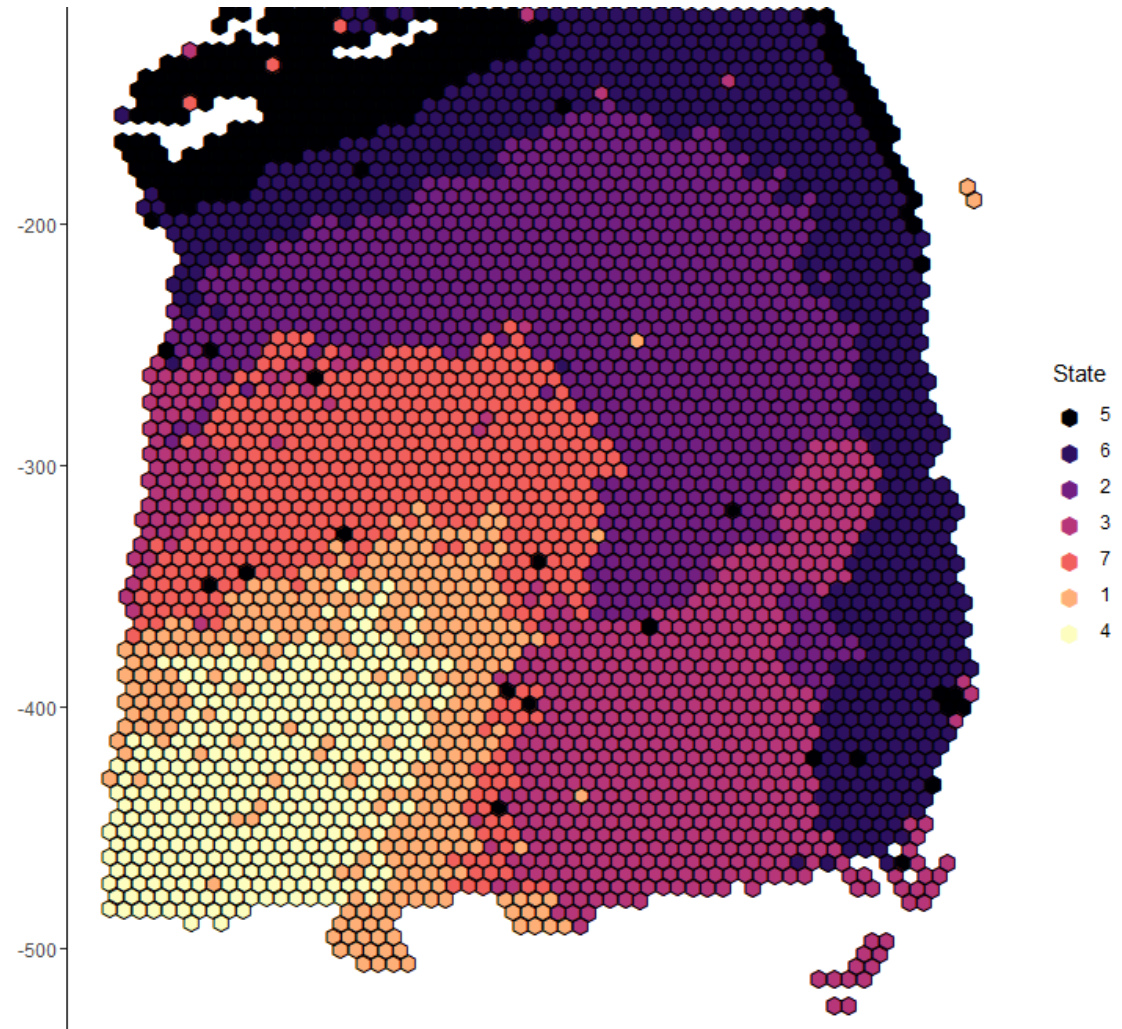
- 7 clusters
- 5 PCs
- Smoothing parameter = 4
- ARI = 0.293





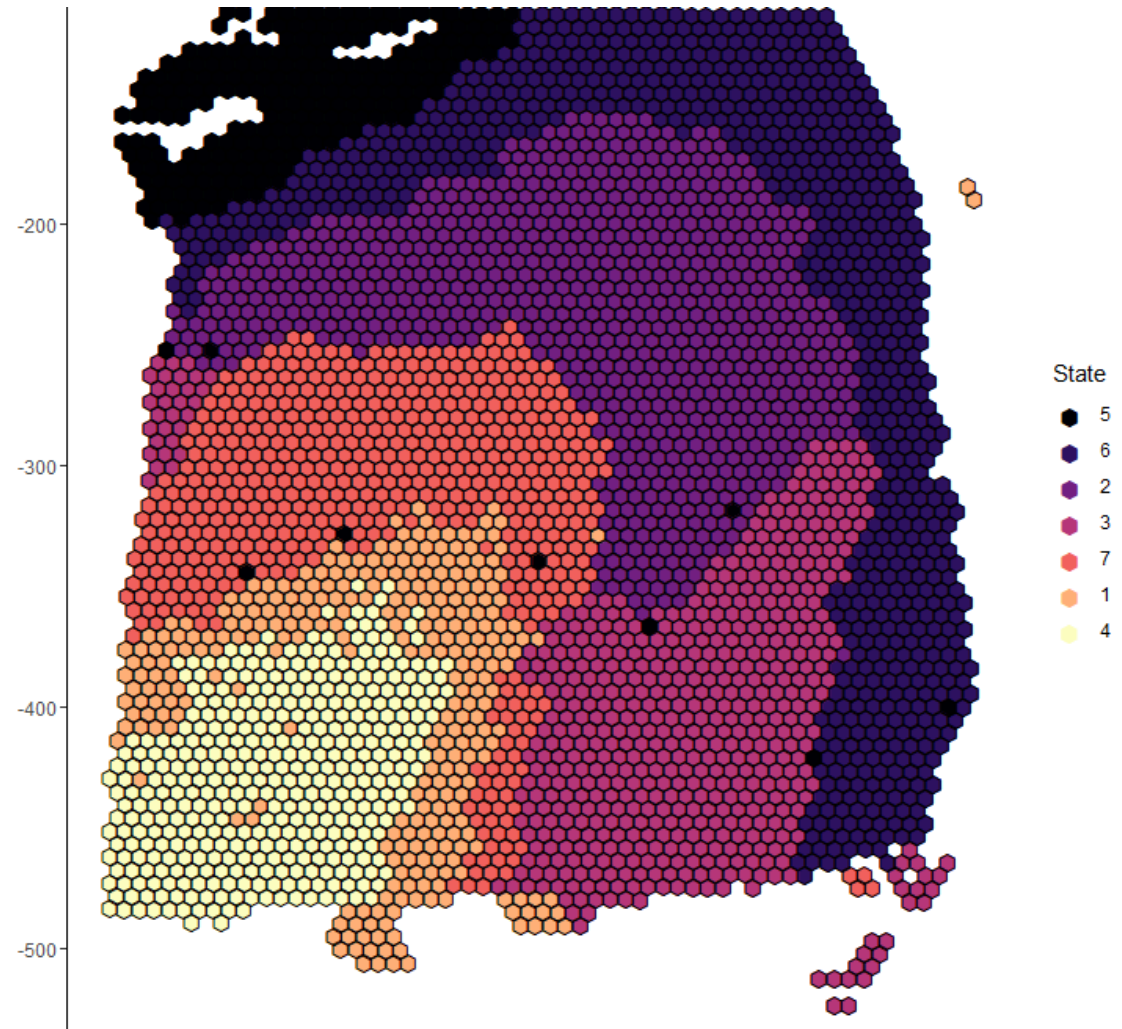
# MCMC Clustering

- 7 clusters
- 9 PCs
- Smoothing parameter = 2
- ARI = 0.299



# MCMC Clustering

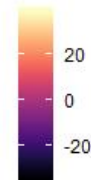
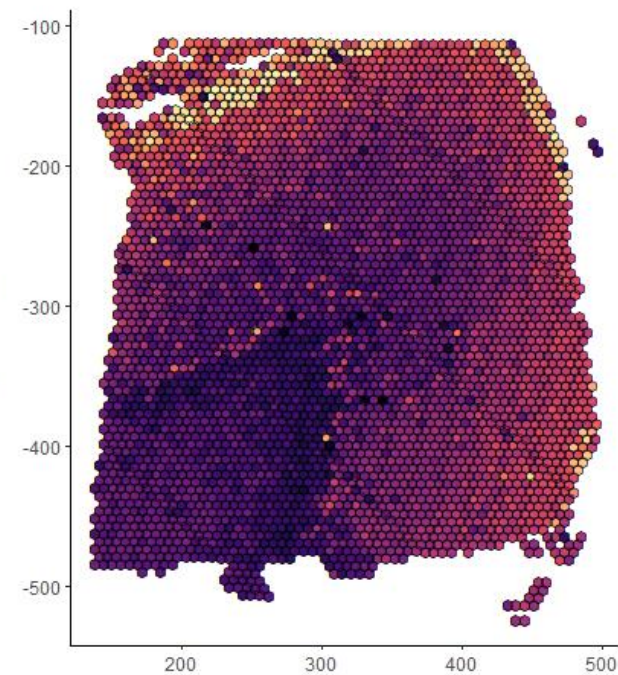
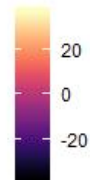
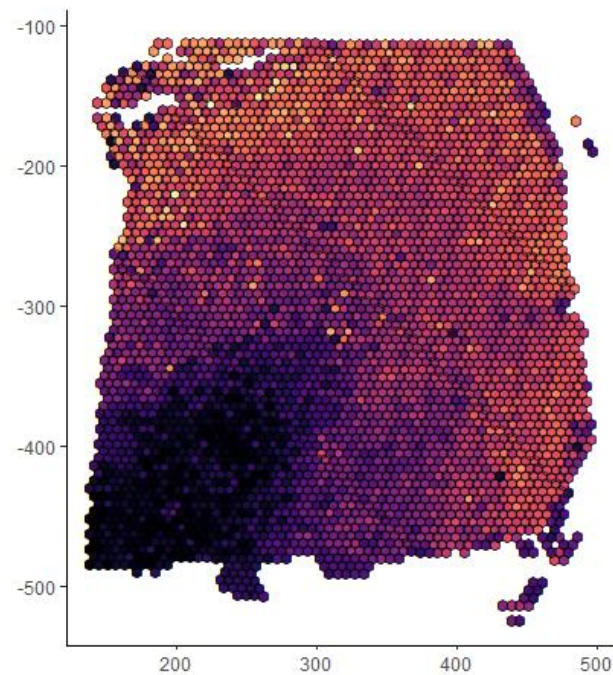
- 7 clusters
- 9 PCs
- Smoothing parameter = 4
- ARI = 0.308



# Other things to try

- Selecting PCs that have a spatial pattern
- Changing number of clusters

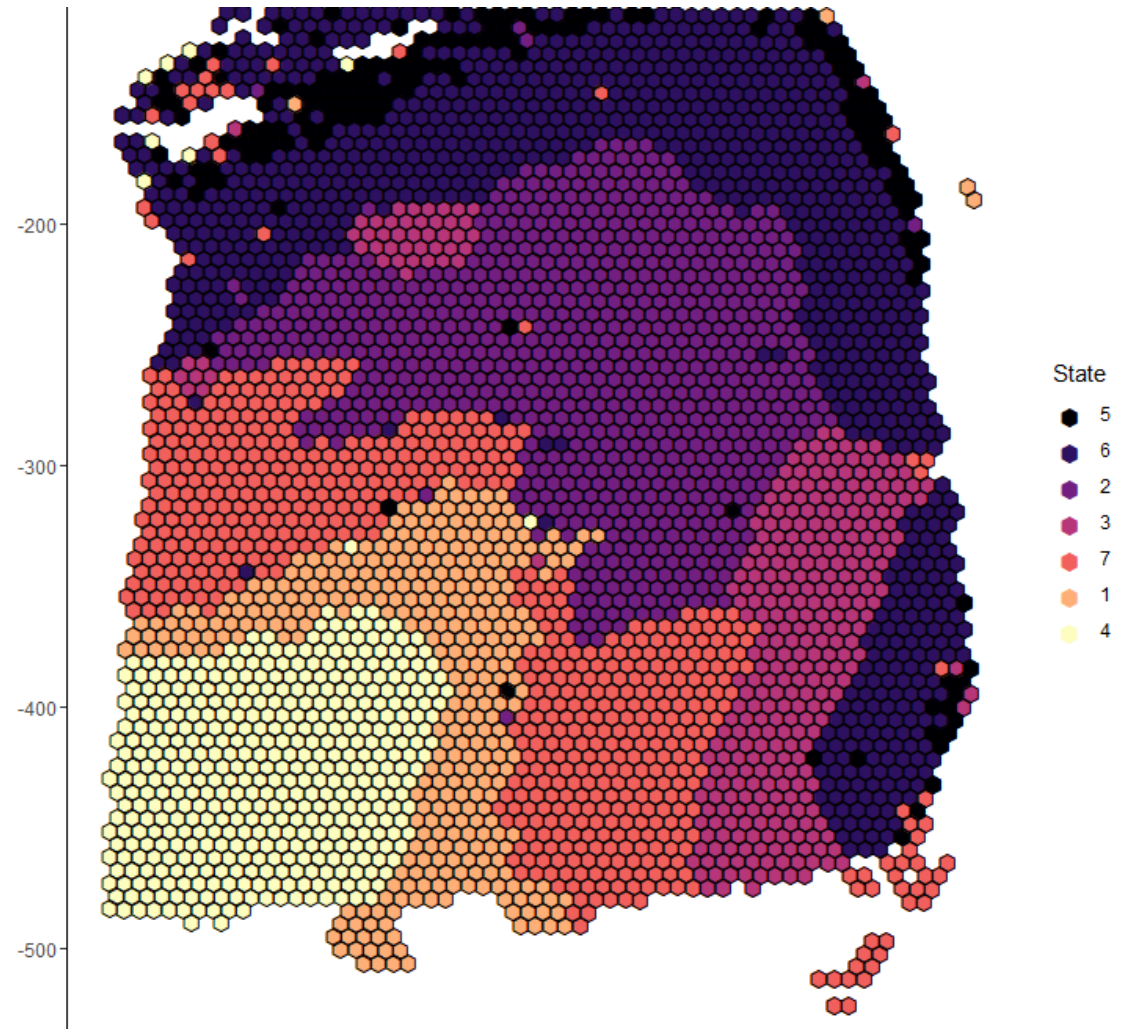
# t-SNE





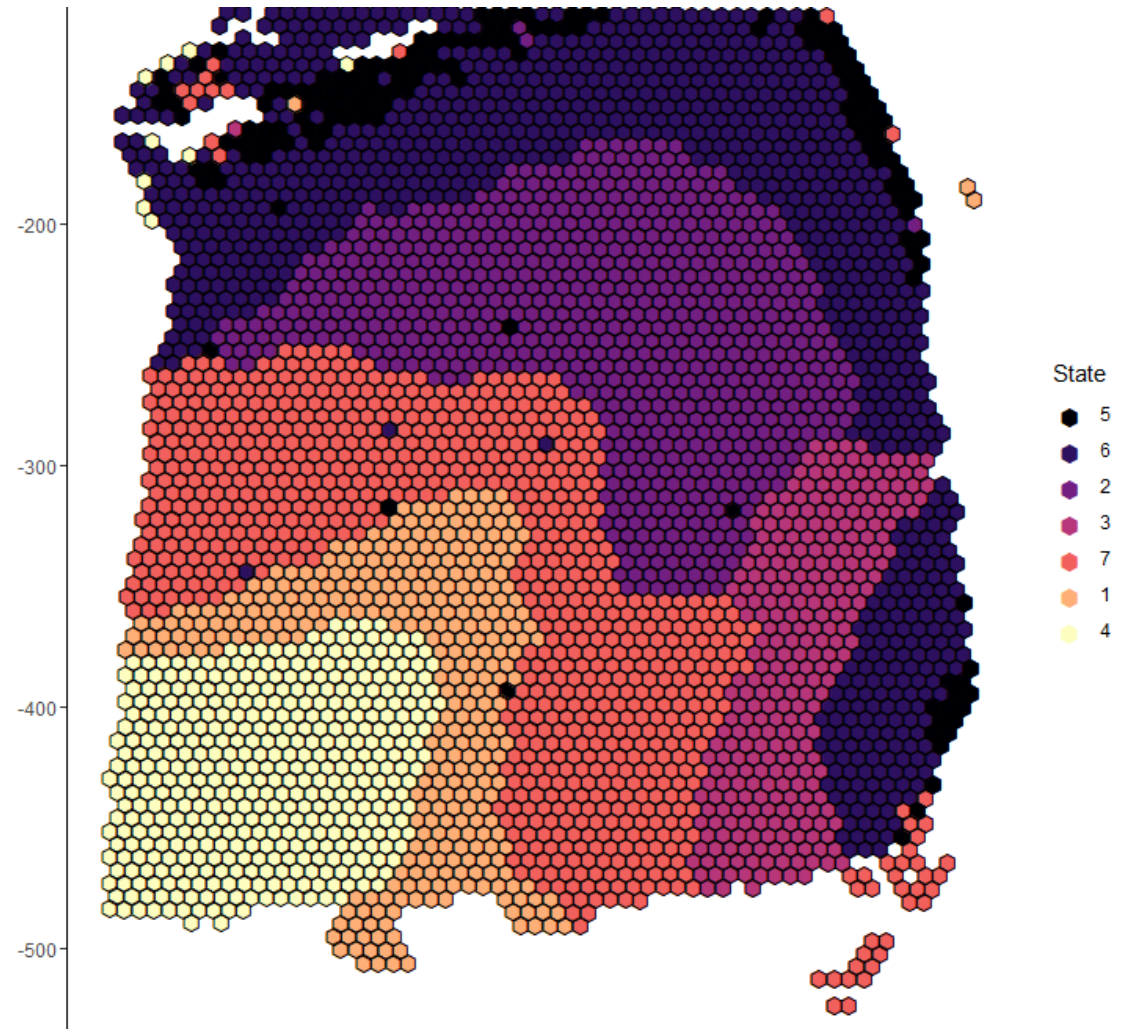
# MCMC Clustering

- 7 clusters
- t-SNE
- Smoothing parameter = 4
- ARI = 0.332



# MCMC Clustering

- 7 clusters
- t-SNE
- Smoothing parameter = 6
- ARI = 0.332



# Other things to try

- Tuning t-SNE perplexity
- UMAP
- Changing number of clusters