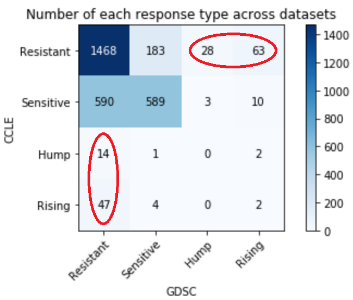
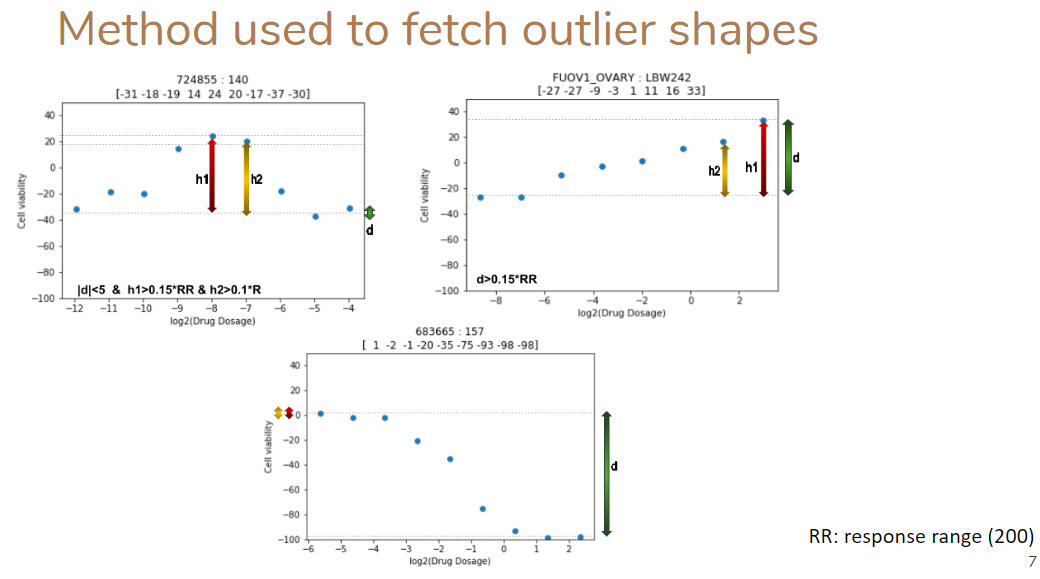
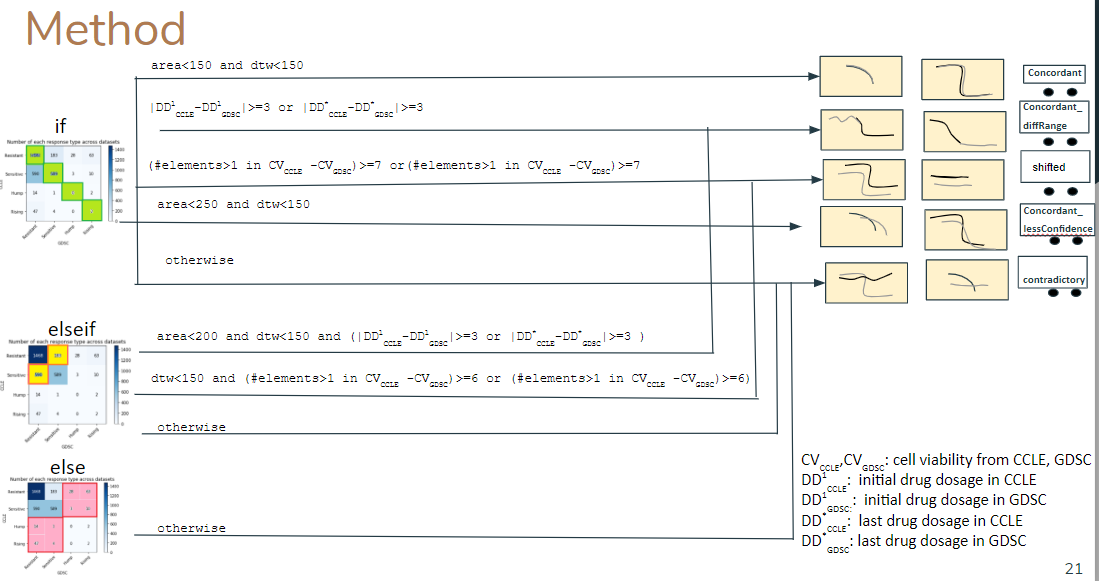
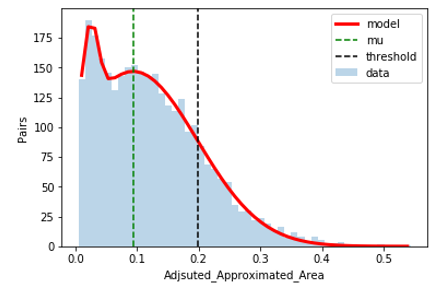
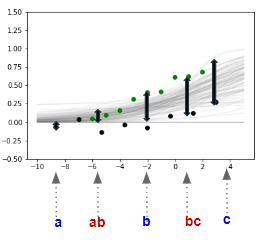
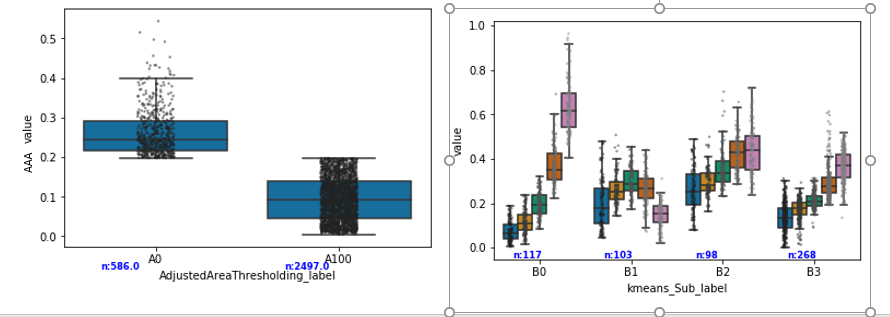
1. Separate cl-dr pairs which are outliers
   * FAILED ATTEMPT: used k-means and agglomerative clustering (using “euclidean” and “pearson correlation” as distance metric between the data points/curves) to separate out outlier cell line-drug pairs and visualized the clustering results using PCA and tSNE
   * Exploited the shape/structure of dose response curves.

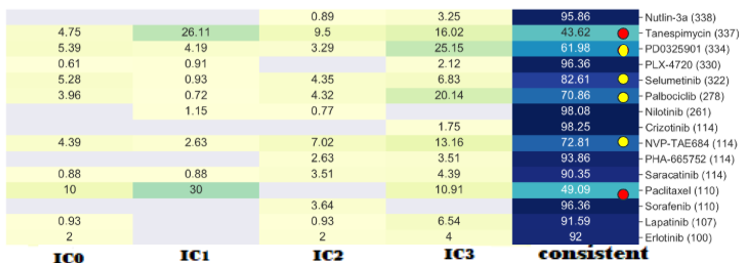
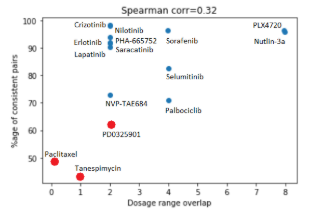


RESULT: The outlier curves is not true behavior of a drug as only 2 common outliers across both datasets.

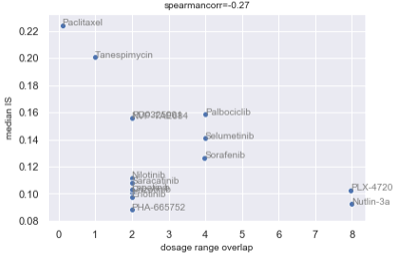
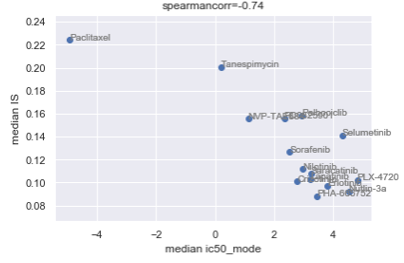
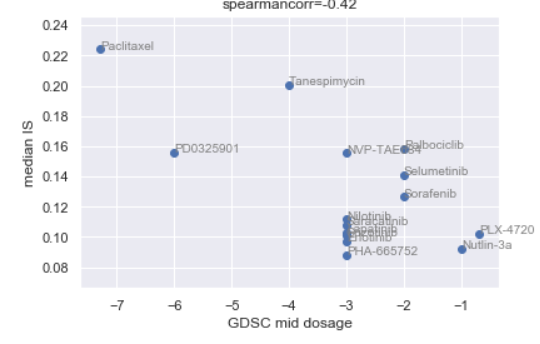
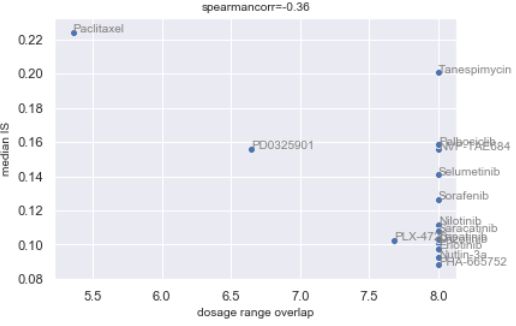
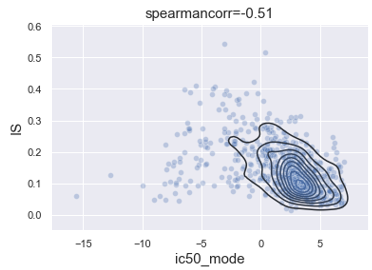
Also, outlier curve in 1 data is Resistant pair in another => **Hump/Rising shaped curve =>** Resistant

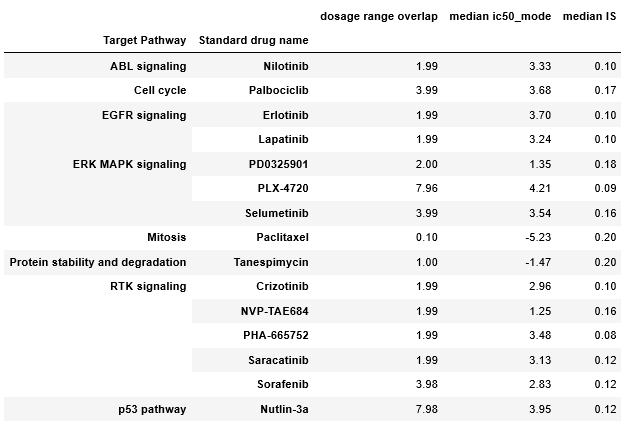
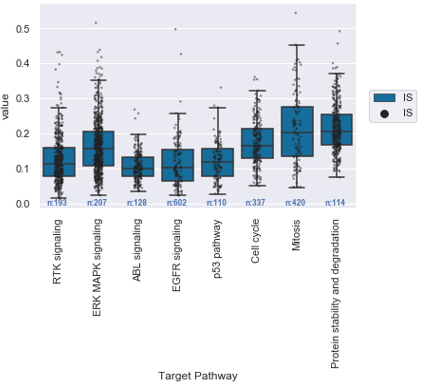
1. Classify Common Pairs
   * A) Using property of curves (Disadv: ths, doesn’t handle different dosage range, doesn’t consider jump/rising as DNW )  
     
   * Fit Bayesian sigmoid curve on union of points/merged dataset. So curve fitting solves dual purpose: 1. Get ic50, m , mae from the fitted curves on union of points 2. Use the features from fitted curves to classify the curves, if needed
   * B) AAA/IS (Adjusted approximated area/Inconsistency score): approximated area between fitted curve using 5 points on hypothetical curves on lower and higher end of HDI.  
     
   * Using curve fitting -> HDI feature extraction -> AAA th and clustering based on 5 f eatures (th free, takes care of outlier curves and diff drug dosage)-> make maps

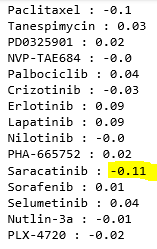
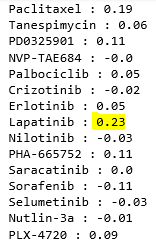
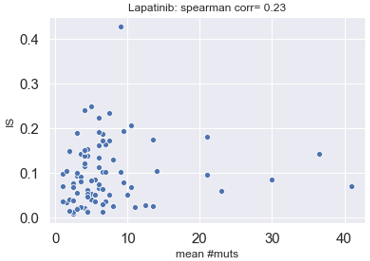
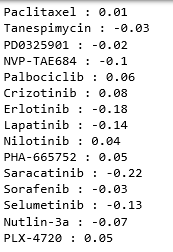


 ­­­­   
  
Faulty drug: Paclitaxel, Tanespimycin , PD0325901, palbociclib, NVP-TAE684, Selumitinib

1. Show same trend but avoiding the usage of any method where thresholding is needed

* **DRUG level analysis (only on sensitive CLs tested for all 15 drugs)**  
     
  (dosage range not on log scale)  
     
  



* Properties common in v bad drugs: low ic50
* Properties common in good drugs: some drugs have similar target pathway
* **CL level analysis**
  + Checked whether jaccard similarity correlates with IS (for all genes and CDG)  
    Drug-wise sp corr using CDG  
    
  + Checked If mean number of mutations correlate with IS (for all genes and CDG)  
    Drug-wise sp corr using CDG
  +  
  + No pattern in IS vs GE profile similarities (top 1k highly dispersed genes)  
    Per drug correlation:
  + 
  + No pattern in IS vs med GE (avged across ccele and gdsc)
  + Per drug correlation:  
    