Was als nächstes ansteht:

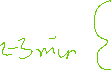
* Zusätzliche gene sets (pathways) auswählen
* Biotypes bestimmen für Datenreduzierung, dann Varianz-Threshold, constant genes?
* PowerPoint Themen aufteilen

Projekt Proposal:



Gliederung:

1. Hallmarks of cancer (Hanahan, Weinberg 2011) Lotti
   1. Biologischer Hintergrund



* 1. Tumor metabolism
  2. Prostate adenocarcinome

1. Welche Daten haben wir? Carla
   1. Genesets + additional genesets



* + 1. Überlauppungen visualisieren
  1. TCGA Expressionsdaten + clinical annots 🡪 preprocessing
     1. Histogramme von mean/sd/var



* + 1. Clinical annots eventuell, -> verteilung merkmale
  1. Vergleich PRAD v/ tumor
     1. Boxplots: means, Var oder SD
     2. Zwischen 5 subtypen
     3. Zwischen tumor / normal samples

1. Welche Fragen wollen wir beantworten? / Welche Ziele haben wir? Wie wollen wir unserer Fragen beantworten? Laura, Fabi
   1. Was für cancer hallmark pathways gibt es?
   2. how many shared genes are there between these gene sets?



* 1. Are there patterns and clusters of tumor samples?
  2. Differences tumor cells vs. Normal cells
  3. Can we predict a regression modell
  4. Usw.

1. Bis wann wollen wir was erledigen 🡪 Timeline



* 1. Ca. 8 Wochen
  2. Preprocessing bis wann
  3. Fertig 1 woche vorher
  4. Wann Bericht anfangen?
* Was sind die Erwartungen ans Project Proposal? Code? Nicht unbedingt
  + Literatur:
    - „summary of literature on this dataset”? Schwerpunkt PRAD
  + Questions we want to address?
    - Was für cancer hallmark pathways gibt es?
    - how many shared genes are there between these gene sets?
    - Are there patterns and clusters of tumor samples?
    - Differences tumor cells vs. Normal cells
    - Can we predict a regression modell
  + Timetable: Ziel -> 1 Woche vor final presentation mit Bericht fertig sein