* Das Inhaltverzeichnis ist richtig nice; das hatte bisher noch niemand☺
* kA, ob das funktioniert, aber es wäre optisch sehr schön, wenn das HTML in Blocksatz wäre: <https://stackoverflow.com/questions/43222169/how-to-justify-the-text-to-both-sides-when-knitting-html-in-rmarkdown>
* While in the pase decades breast cancer therapy 🡪 past decades
* Ich kann es faktisch nicht begründen, aber gefühlt hätte ich ohne Komma gesagt: One promising method is the computerbased approach, with possible
* Driver mutations (DV) are what is to be seen as the ticket of admission to the different stages of cancer growth. 🡪 sounds not so scientific
* k-ras and b-raf encoding kinases 🡪 both encoding kinases
* On the other hand the most illustrious tumour suppressor 🡪 On the other hand, 🡪 hier hätte ich auch ein Komma gesetzt
* several parameters 🡪 be precise which parameters
* ACHTUNG:  #CITATION A promising approach is targeting so-called Second Site Targets (SSTs), which are gene mutations interacting genetically with driver mutations to increase cell viability. # CITATION If knocking out those genes leads to cell lethality, it may lead to new treatment options. 🡪 hier habt ihr noch #Citation
* so-called Second Site Targets (SSTs) 🡪 die Abkürzung habt ihr vorher schon mal eingeführt
* Die library importation könnt ihr von mir aus drinnen lassen; aber das install.packages 🡪 die Section würde ich ganz einen einen eigenen Codejunk geben, den ihr dann einfach ausblendet, weil sonst wird das super viel 🡪 bei den Libraries kann man argumenbtieren, ob man die drinnen lässt, weil ihr viele spezielle verwendet; die Entscheidung überlasse ich auch; also von meiner Seite her könnt ihr auch die ausblenden in einem eigenen Codejunk; sodass die erste sichtbare Zeile #Setting a sys-path wäre
* *# installation of required packages 🡪 schaut auf einheitliche Kommentare; weil danach kommt mal alles in so einem Format: #TEXT und nicht wie hier # text*
* To get a feeling for the dataset we will have a look at the given data structure. 🡪 … we had a look at ….
* Afterwards,the dataframe could be subsetted 🡪 zwischen afterward , und the fehlt euch ein Leerzeichen
* The next step was to do some experiments,  🡪 experiments würde ich durch z.B. broad data analysis ersetzten oder sowas; weil das Wort ist so lala beim Lesefluss
* *# expanding the colourpalette to 28 🡪 dieser Kommentar beginnt auf eineml klein und mit Leerzeichen zwischen # und Expanding*
* An issue was the very low level of expression of most genes, seen in the graph above 🡪 wenn ihr schon so schöne Figure X macht; dann sagt doch ruhig ….seen in Figure2
* In this phase of the project is was open if a  🡪 is sollte, denke ich mal, ein it sein
* Next, boxplots of expression values for driver mutations suggested  🡪 mutations suggest
* Ich weiß nicht wie es euch geht; aber ich zähle 12: Fig. 3 Expression values 15 selected driver mutations
* *#determining optimal cluster number k 🡪 Determine*
* Plotting wss function und using ggplot to plot output 🡪 hier Kommentare mit kleinem p und Leerzeichen; gleiches für selecting only genes mutated at least once and with Ceres <=0 🡪 generell das mit den Kommentaren checken
* Bei eurem optimal cluster-plot da könnt ihr ruhig die Annotation mit den Linien (also BerastCancer-1 (orange), Berast-Cancer-2 (rot) etc. was da halt unter dem Plot ist weglassen; das geht mit legend.position = „none“; weil das finde ich ist eh klar, dass man da die Sample plottet
* 2 wss plots fände ich nicht nötig; sucht euch einen aus ☺
* Ich würde gefühlt nach thus ein Komma setzten: Thus all potential SSTs candidates were expected
* While the ERBB2 driver mutation was found to have the lowest essentiallity  🡪 is found; oder?
* ?????????:  I created exaclty the same matrix little bit further up in the code. Perhaps there is no need to recreate a new dataframe here? in my perspective it is better to have less code.. i already discarded most of the exploratory stuff.
* Kommentare bei Preperation of the data and Wilcox performance nicht einheitlich zu vorhin
* Genau; wie ihr schon auf WhatsApp angemerkt habt; bitte dringend nicht alles plotten; aber das macht ihr schon☺ 🡪 ich kenne paged\_table nicht; aber würde das mit head funktionieren???
* Im code junk: *#get summary of distribution; wieder alle Kommentare cheken; auch die auf der Seite*
* Kommentare bei SST\_Data; wieder checken
* generated by the second lapplyloop. 🡪 in eurem code sehe ich da einen sapply?
* For the regression model, we want to predict the CERES score 🡪 wanted
* To run multilinear regression,  🡪 to run a multilin. Reg.
* Auch bei der Regression wieder die Kommentare checken
* To run regression 🡪 to run a regression
* Geth das was da bei run gresion kommt als Aufzählung; weil so finde ich das nicht so schön, um ehrlich zu sein, mit –/+ mal mitten drin
* Achtung; eure Heatmaps heißen alle Figure 8: QQ plot XXXX 🡪 das sind aber heatmaps; oder☺?
* Und wie ihr sagt; wäre cool, wenn ihr diese ganzen Plots die da so kommen (diese 8) evt. In ein schönes Pannel geben könntet; das kann man Händisch in Photshop oder kA was vorbereiten und dann das panel als Foto importieren und einfach den Code halt den Plot nicht printen lassen, sondern nur produzieren
* Die ganzen Tabellen zu eurern Plots da unten bei den QQplots müsst ihr nicht unbedingt machen; also weil das macht es nur etwas verwirrend, finde ich persönlich
* To run the Multilinear Regression, 🡪 wenn ich das richtig im Kopf habe, habt ihr bisher multilin reg. Immer klein geschrieben
* CERES scores (x1 … xn; indep.variables) Followed 🡪 followed gehört klein
* runs the test Data 🡪 data klein
* For the Spearman correlation, the rank-correlation coefficient can take values from +1 to -1 + rho ~ 1: perfect positive association of ranks + rho ~ 0: no association of ranks + rho ~ -1: perfect negative association of ranks 🡪 da fhelt der . am Ende
* To further interprete the p-value, we construct a H0 / H1 hypothesis:  🡪 auch hier; eine Aufzählung möglich???; weil so haben wir wieder eure Aufzählungszeichen mitten im Text
* How about doing a linear regression model and comparing the F-statistic values? # Discussion of sample size In summary it can thus be argued that these multilinear regression models are not effective in predicting the impact of driver mutations on cell viability. One possible reason for unfitting regression models could be the small sample size of only 28 events (breast cancer cell lines). A common rule of thumb for building a regression model is to calculate the limiting sample size: 🡪 ist das wieder ein noch fehlender Kommentar???
* If this equation is fulfilled, the model is supposed to be reliable.(Harrell 2001).  🡪 vor(Harrell, 2001) ist ein . und in der Klammer fehlt das Komma zwischen Autor und Jahr
* otherwise with the entry for a given cell line being âNAâ. Further data cleanup was performed by removing all genes which contained all âNAâ values and  🡪 ACHTUNG: Formatierungsfehler mit den Kästchen
* With this findings 🡪 these
* Overall, it must however be noted that although 🡪  Overall, it must be noted that although
* The final step in this project was to exploit 🡪 nach de final step in the project würde ich ein Komma setzten
* on small training sets.(Banko and Brill 2001). –> wieder . vo () und gehört eben nicht (Autoren, Jahr 🡪 also wie oben angemekrt das Komma?) 🡪 gleiches gilt ggf. für LIbbrecht
* Banko, M. and E. Brill (2001). Scaling to very very large corpora for natural language disambiguation. Proceedings of the 39th Annual Meeting on Association for Computational Linguistics. Toulouse, France, Association for Computational Linguistics: 26-33. 🡪 Formatierungsfehler; dieser ist zwei Punkte eingerückt
* Sehr schöner Bericht; sehr gut gemacht!