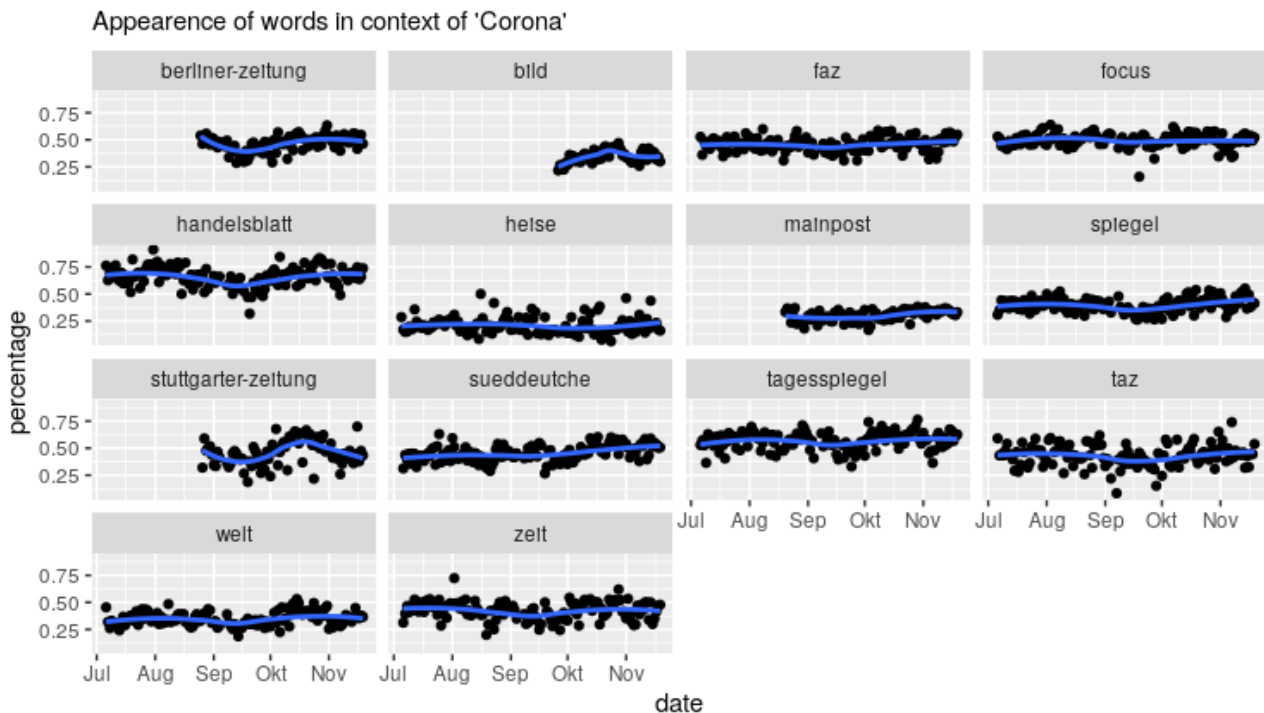


Als erstes wurden aus dem Datensatz alle Artikel rausgefiltert die Wörter wie Corona, Sars-CoV-2 oder Covid-19 beinhalteten, ohne Rücksicht auf Groß- und Kleinschreibung. Die Datenpunkte ist der prozentuale Anteil der Artikel einer Zeitung an einem Tage, der eben genau diese Wörter enthält. Der Datensatz umfasst ca. 61.000 Artikel. Die Statistik dazu sieht wie folgt aus:

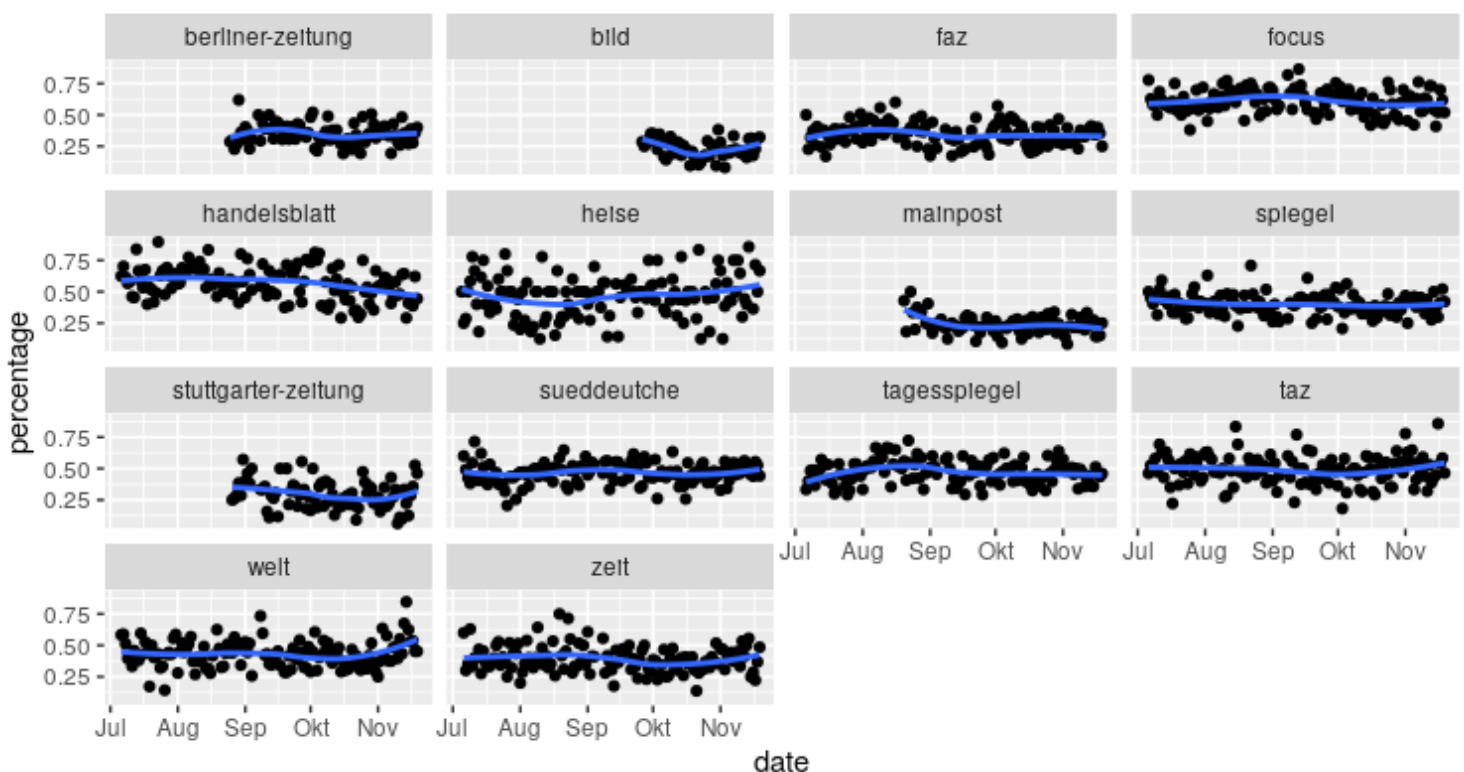


Dieser neu bearbeitete Datensatz (nenen wir ihn „Corona-Datensatz“ wurde nun als Grundlage für folgende Untersuchungen genommen:

1. Verharmlosung:

Hier wurde wieder der prozentuale Anteil der Artikel aus dem Corona-Datensatz rausgefiltert die im Title oder Text diese Wörter enthielten: "harmlos|sinken|gesunken|sicher|unbedenklich|ungefährlich|nicht.ansteckend|gefahrlos|risikolos|nicht.risikant" (Dieser String ist ein regulärer Ausdruck, der in der Funktion grepl() zur Untersuchung von Character Vektoren in R benutzt wurde. Der senkrechte Strich stellt ein logisches oder dar.)

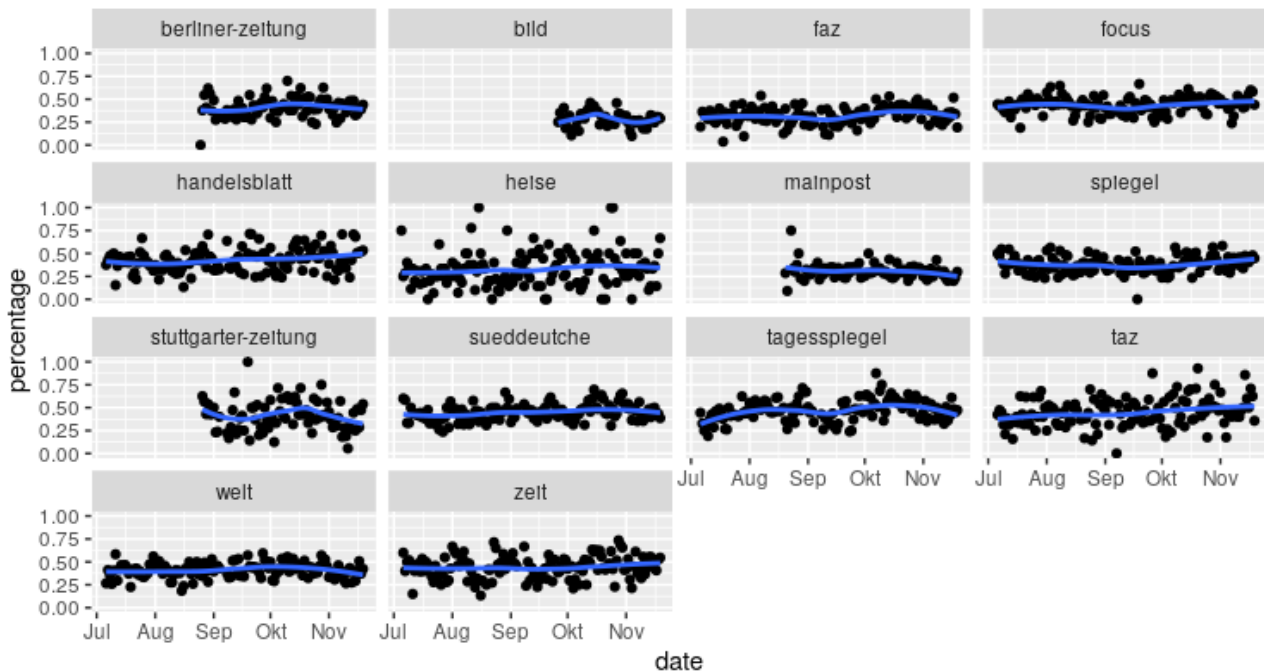
The appearance of synonyms for the word 'harmless'



2. Regeln:

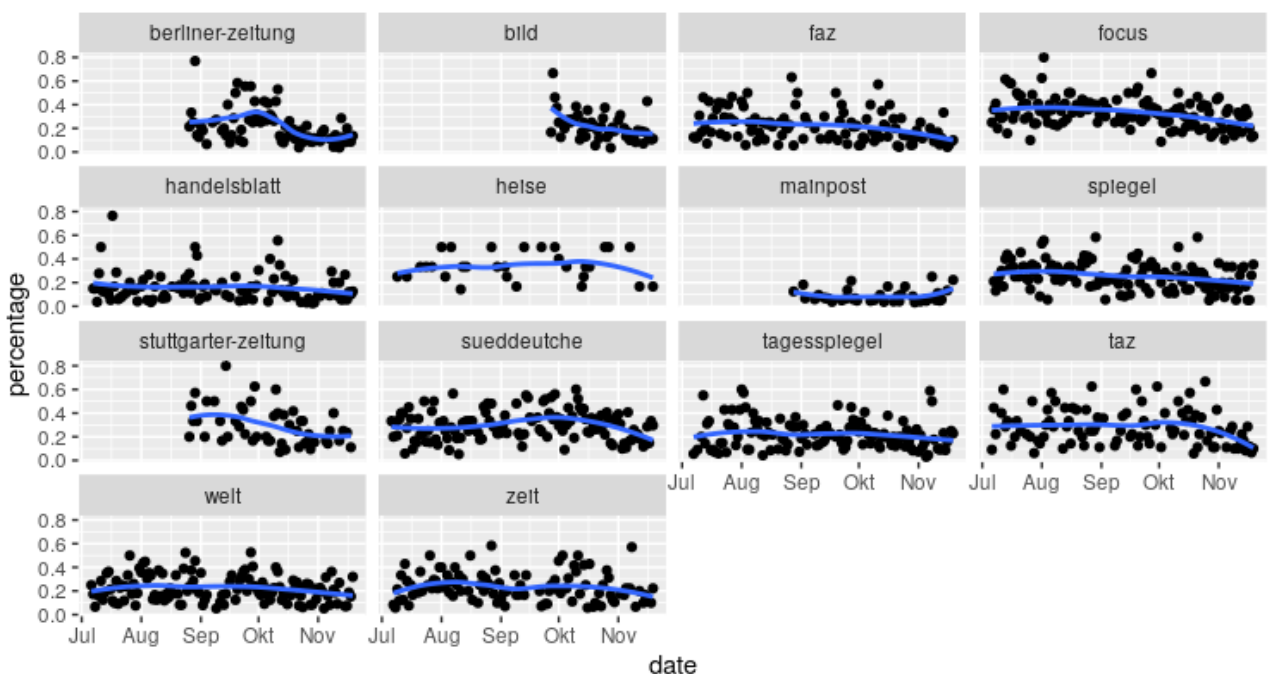
Als nächstes wurde untersucht in wie weit, in dem Corona-Datensatz, die Regeln und Restriktionen vorkamen. Dabei wurden wieder diese Wörter zu Filterung genutzt: "regeln|beschränkungen|beschränkung|restriktion|einschränkung|richtlinien|verordnung|kontaktbeschränkungen|maßnahmen". Dabei entstand diese Statistik:

Appearance of words and phrases which are about Restrictions



Der gefilterte Corona Datensatz im Bezug auf die Regelungen umfasste ca 25.000 Artikel. Dieser neue Datensatz wurde nun noch einmal mit diesen Wörtern gefiltert: "verstoß|nicht.eingehalten|party|unzufrieden|strafe|bußgeld", um zu schauen ob eventuell, die leute sich mehr oder weniger an die Regeln halten. Die Datenpunkte der folgenden Statistik sind nun wieder rum der prozentuale Anteil der Artikel- an einem Tag, von einer Zeitung- die sich mit Corona-Beschränkungen befassen:

Appearance of words and phrases which are against Corona restrictions



3. Querdenken:

Als letztes wurde die selbe Methode auf das Vorkommen von Wörtern im Kontext der Corona-Leugner untersucht. Der Suchstring hierfür: "qanon|querdenken|ideologie|lügenpresse|afd|maulkorb|rechte.szene|rechtsextrem|verschwörung|leugnen|leugner". Die daraus entstandene Statistik sieht wie folgt aus:

Appearance of words in context of denying the Corona pandemic

