

GKL671

@author Kacper Bohaczyk

@version 20-03-2023

- **Data Science** Data Science ist ein Fachgebiet welches Wissen aus verschiedenen Bereichen wie Statistik, Mathematik und Informatik extrahiert um sie zu analysieren und Muster bzw. Zusammenhänge zu identifizieren

- **Machine Learning**

Machine Learning ist ein Bereich der Künstlichen Intelligenz, der sich auf die Entwicklung von Algorithmen und Modellen konzentriert. Dadurch wird dem Computer ermöglicht Vorhersagen und Entscheidungen zu treffen.

- **Künstliche Intelligenz**

Künstliche Intelligenz bezeichnet Systeme, die in der Lage sind, menschenähnliche Intelligenz auf verschiedenen Ebenen zu zeigen. Gute Beispiele können die Erkennung von Bildern, das Verständnis natürlicher Sprache sowie die Entscheidungsfindung.

- **Data Science**

Data Science ist ein interdisziplinäres Fachgebiet, das sich mit der Extraktion, Analyse und Interpretation von Daten beschäftigt, um daraus Erkenntnisse und Wissen zu gewinnen. Es vereint verschiedene Aspekte der Statistik, Mathematik, Informatik und domänenspezifischen Kenntnisse, um Werte aus Daten zu extrahieren.

- **Data Scientist**

Ein Data Scientist ist ein Experte auf dem Gebiet der Datenwissenschaft, der Techniken und Methoden verwendet, um Daten zu sammeln, zu verarbeiten und zu analysieren, um Erkenntnisse und Wissen zu gewinnen. Data Scientists kombinieren verschiedene Fähigkeiten und Kenntnisse in den Bereichen Statistik, Mathematik, Informatik, Datenanalyse und

Geschäftsstrategie, um komplexe Datenprobleme zu lösen und Erkenntnisse aus großen Datenmengen zu gewinnen.

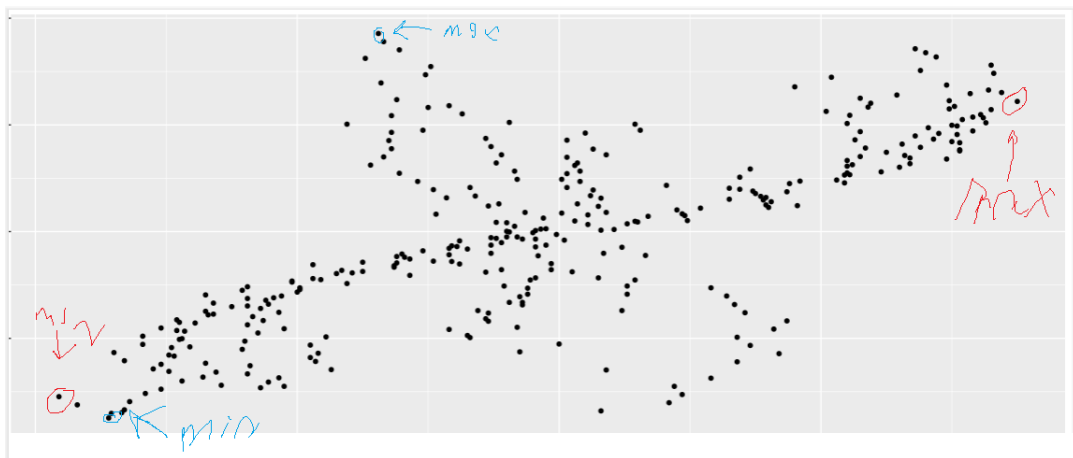
- **Datenexploration**

Die Datenexploration ist ein Prozess der Untersuchung eines Datensatzes, um Muster und Zusammenhänge zu identifizieren. Es werden verschiedene Techniken wie Datenbereinigung, Datenvisualisierung und statistische Analyse.

- **Clustering**

ist eine Methode des unüberwachten Lernens, die es ermöglicht, ähnliche Datenpunkte in Gruppen zu unterteilen. Ziel ist es, natürliche Gruppen und Cluster von Datenpunkten zu finden, die sich ähnlich verhalten.

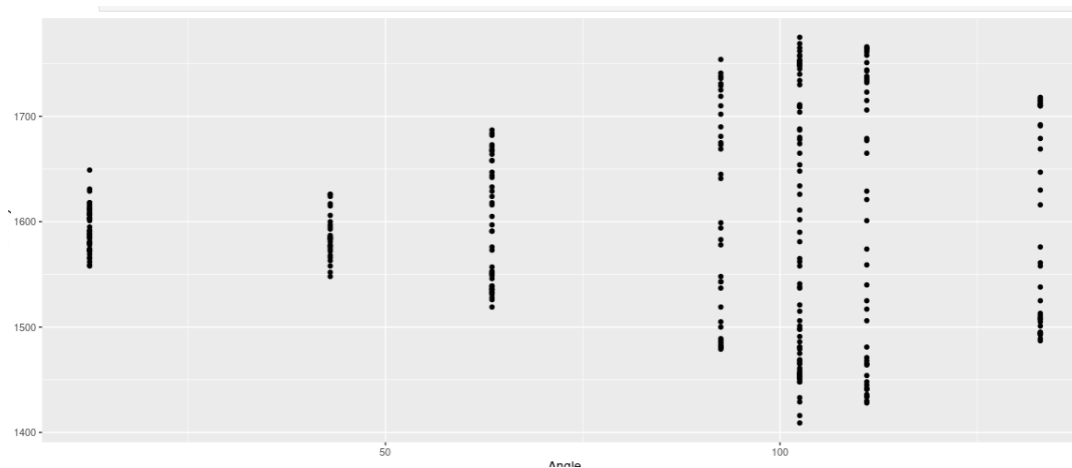
- Ermittle die größte und kleinste Geschwindigkeit in der beobachteten Galaxie und ihre minimale und maximale radiale Position.



Rot: Maximale und minimale Geschwindigkeit

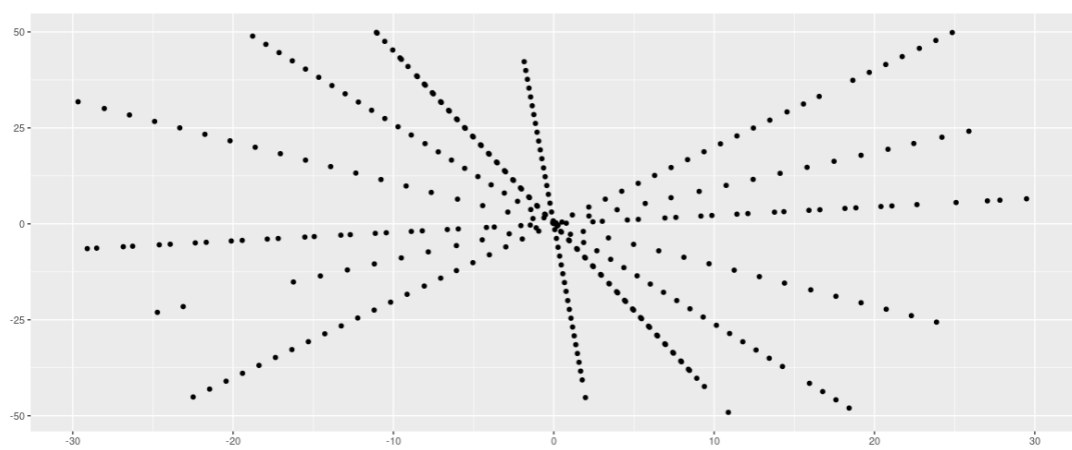
Blau: Maximale und minimale radiale Position

- Ermittle, bei welchem Winkel die extremsten Geschwindigkeiten angenommen werden.



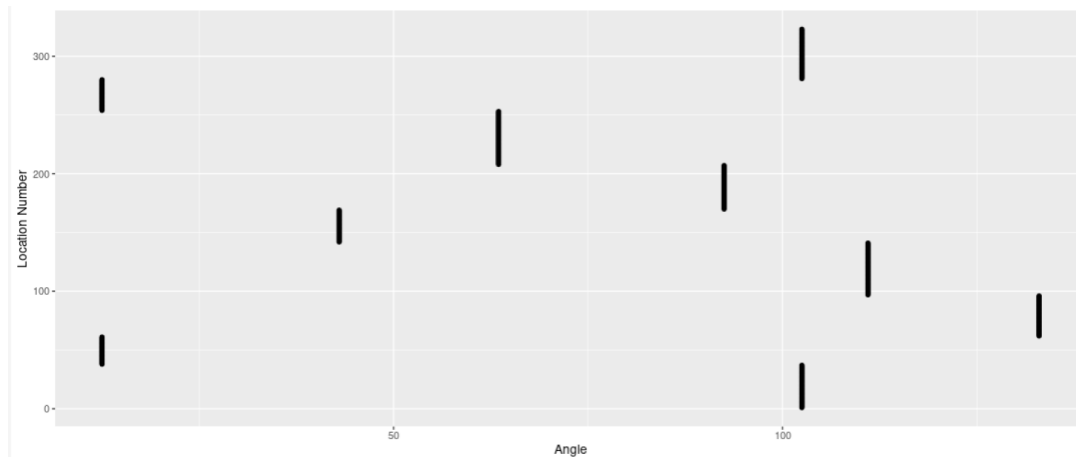
Die Maximalgeschwindigkeit ist ca. bei 105°

- Beschreibe, welche Information sich in einer Punktwolke darstellen lässt, wenn die Nord-Süd-Richtung gegen die Ost-West-Richtung aufgetragen wird.

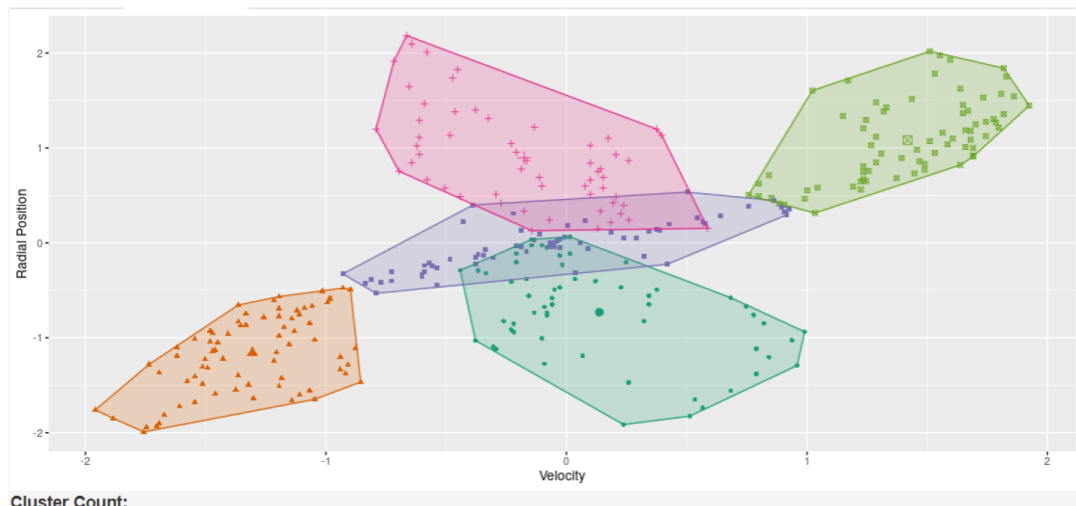


Bei Kreuzung der Ost-West-Position mit der Süd-Nord-Position ist das Diagramm im Verhältnis 1:1. Da können die Winkel der einzelnen Messungen visualisiert werden.

- Überlege, welche Information sichtbar wird, wenn der Winkel gegen die Location Number aufgetragen wird.



Bei kreuzung der Positionsnummer mit dem Winkell kann man sehen wie oft das Teleskop die Galaxie durchquert und Messungen vornimmt, sowie die Reihenfolge der Messungen.



Die Anpassung der Cluster zielt darauf ab, ähnliche Daten in Gruppen zu gruppieren, um eine bessere Veranschaulichung zu ermöglichen. Bei der Entscheidung, welche Cluster angewendet werden sollen, wird dieses Kriterium berücksichtigt. In diesem Beispiel wurden fünf Cluster ausgewählt, um die langsamsten, schnellsten, höchsten und niedrigsten Punkte leicht zu bestimmen.