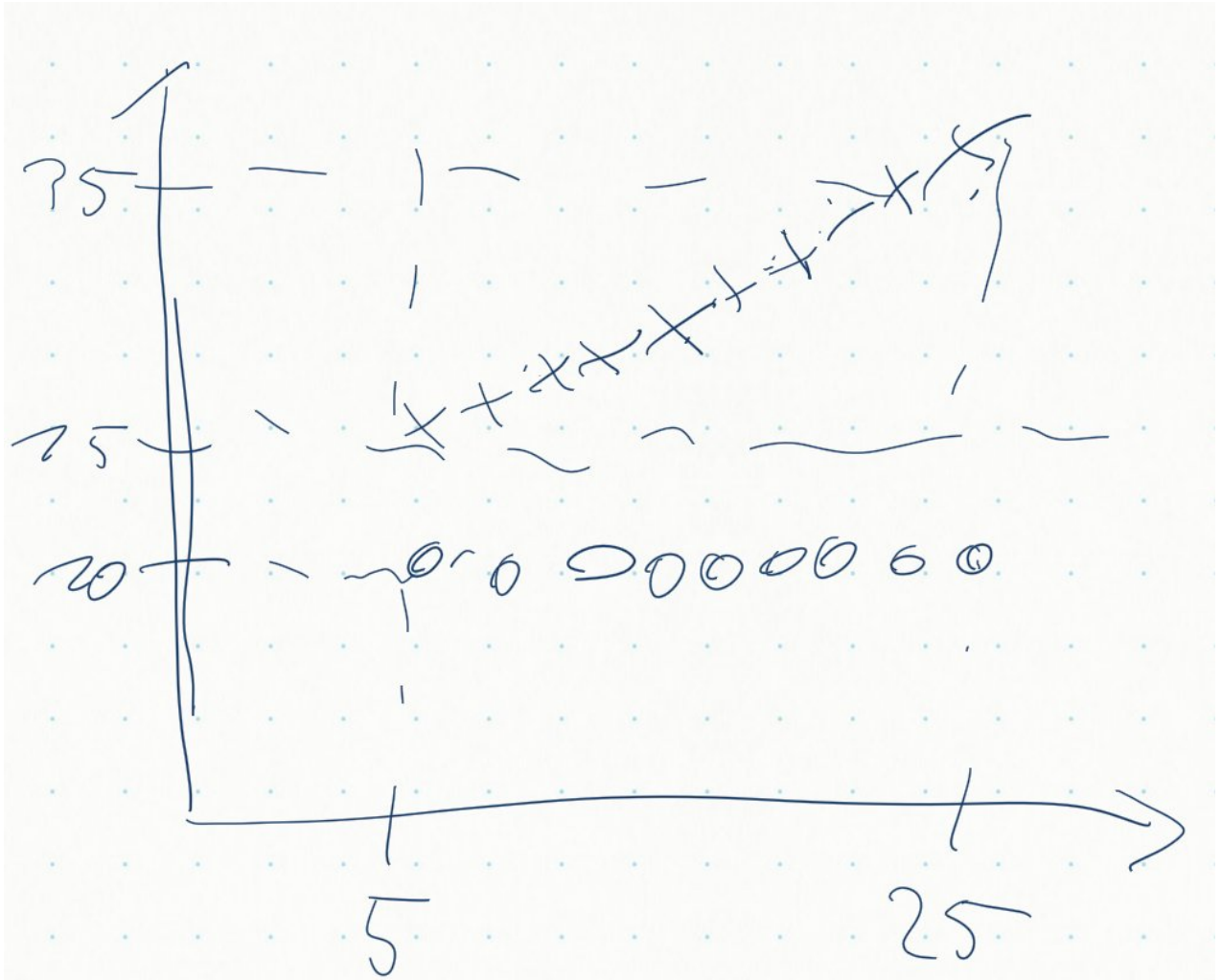


DST-Klausur vom 21.02.2019

Allgemein: Folien dienen eher der Orientierung

Aufgabe 1: Lineare Regression

- Was ist Lineare Regression und was macht Sie? (3 P)
- Wie funktioniert Lineare Regression? welcher Fehler wird minimiert
- Gegeben folgender Scatterplot:



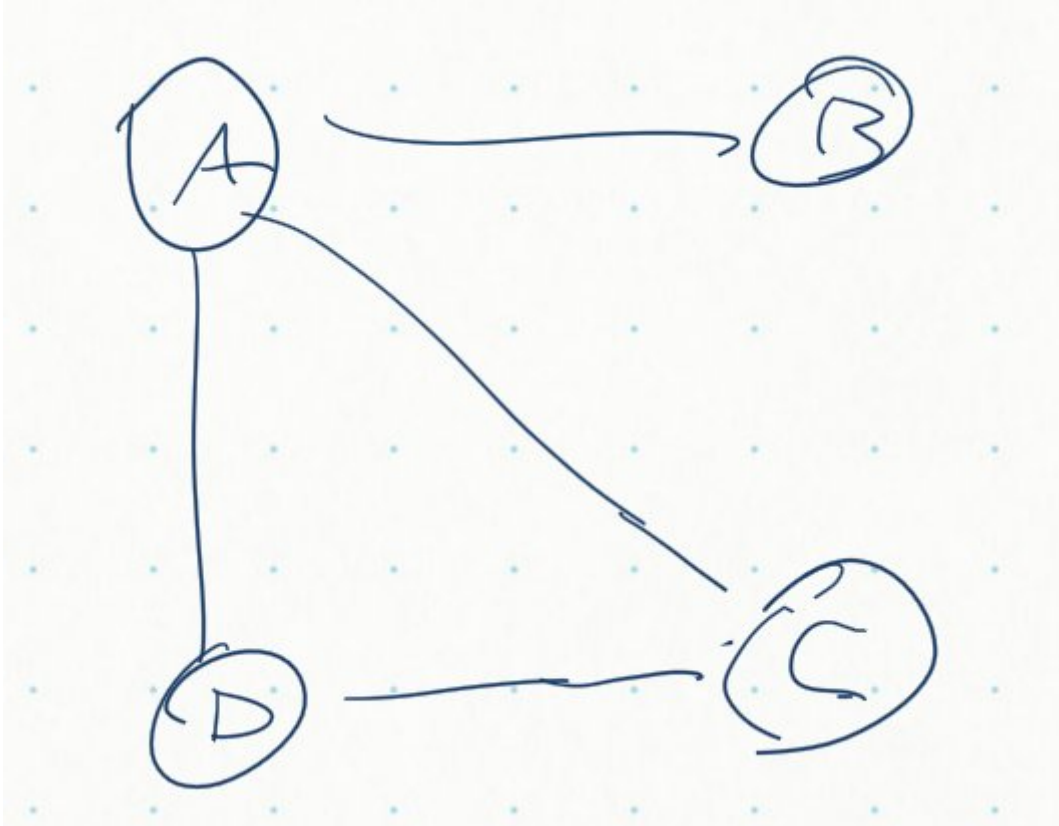
- Wie lauten die Koeffizienten (alles ohne x ausfüllen)?

Variablen	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5
Intercept					
x1	x		x		
x2	x	x			
x1*x2	x	x	x	x	

- Was ist Mean-Clustering? Wie wirkt sich Mean Clustering auf Modell 5 aus?

Aufgabe 2: Social Network Analysis

- Was ist Social Network Analysis?
- Nenne 4 Arten von Knotenanalysen
- Gegeben sei Graph G, Berechne Closeness Centrality für A – D, Diameter, Clustering für A



- Graph:
- Male k-Means-Schritte anhand von gegebenem Scatterplot auf

Aufgabe 3: Multiple Choice

- 8 Fragen à 3 Punkte (tw. sehr folienspezifisch) zu allen Themengebieten

Aufgabe 4: Data Visualization

- Gegeben Dataset {1,2,8,9,12,15,21,24,25,26,30,31,35,87,102} in zufälliger Folge
 - berechne Mean, Median
- Nenne 2 Möglichkeiten, wie das Dataset grafisch dargestellt werden kann (ohne Box-Plot)
- Darstellung malen, die Diskrepanz darstellt
- Berechne folgende Werte:
 - Wie viele Geldeinheiten halten die obersten 20% der Personen?
 - Von wie vielen Personen werden die untersten 33% der Geldeinheiten gehalten?
- Boxplot zu Dataset zeichnen (Tip: 1,5 IQR für Whiskers)
- Nenne und beschreibe 3 Probleme mit Farben in grafischen Darstellungen