

Win-Statistik - Aufgaben für die Kleinklassen w.2Stat-WIN.XX

Ansprechpartner für Korrekturen und Hinweise:

Martin Schnauss, Bledar Fazlija, Ruben Seiberlich, Oliver Bachmann

Quellen

1. Newbold, P., Carlson, W., Thorne, B. (2013). Statistics for Business and Economics (Global Edition). 8. Aufl. Upper Saddle River, N.J.: Pearson Prentice Hall.
2. McKinney, W. (2018). Datenanalyse mit Python. 2. korrigierte Auflage. O'Reilly. ISBN 978-3-96009-080-9. Online in der ZHAW-Bibliothek erhältlich.

Aufgabe 1 (Installation)

10 Punkte

Um auf Windows mit Python arbeiten zu können, laden Sie den Anaconda-Installer (<http://anaconda.com/downloads>) herunter. Wir empfehlen Ihnen, den Installationsanweisungen für Windows zu folgen, die Sie auf der Anaconda-Downloadseite finden. (siehe Abbildung 1) Stellen Sie nun sicher, dass alles korrekt konfiguriert ist. Um ein Kommandozeilenfenster zu öffnen, klicken Sie auf das Startmenü und gehen Sie in die Eingabeaufforderung (auch bekannt als cmd.exe). Versuchen Sie, den Python-Interpreter zu starten, indem Sie Python eintippen. Sie sollten ein ähnliches Fenster sehen wie in Abbildung 2.

Um Jupyter zu starten, führen Sie in einem Terminal den Befehl Jupyter notebook aus. Auf vielen Plattformen öffnet sich Jupyter automatisch in Ihrem Standard-Webbrowser. In Abbildung 3 sehen Sie, wie es in Google Chrome aussieht.

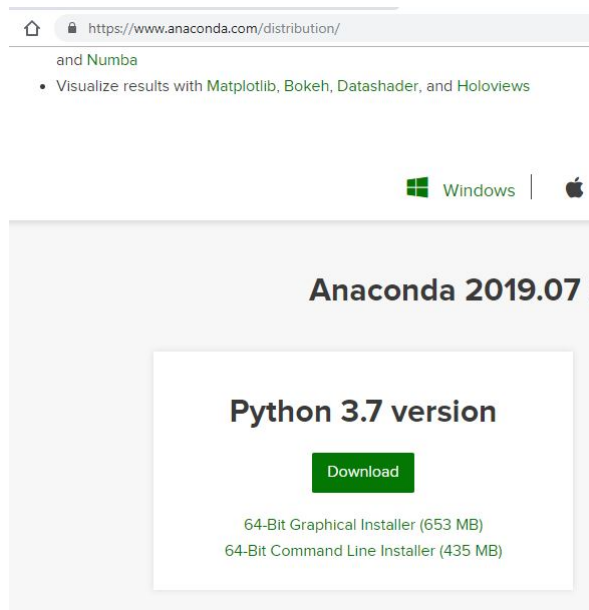


Abbildung 1: Anaconda installieren (auch ältere Versionen sind möglich)

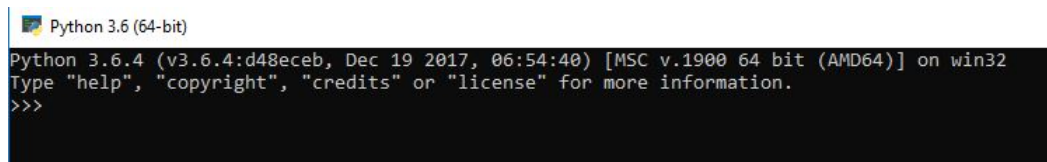


Abbildung 2: Python Interpreter

Bemerkung: Viele Programmierer nutzen Jupyter als lokale Arbeitsumgebung. Es kann aber auch auf Server ausgelagert und aus der Ferne zugegriffen werden. Wir werden darauf nicht

Aufgabe 1 (Fortsetzung)

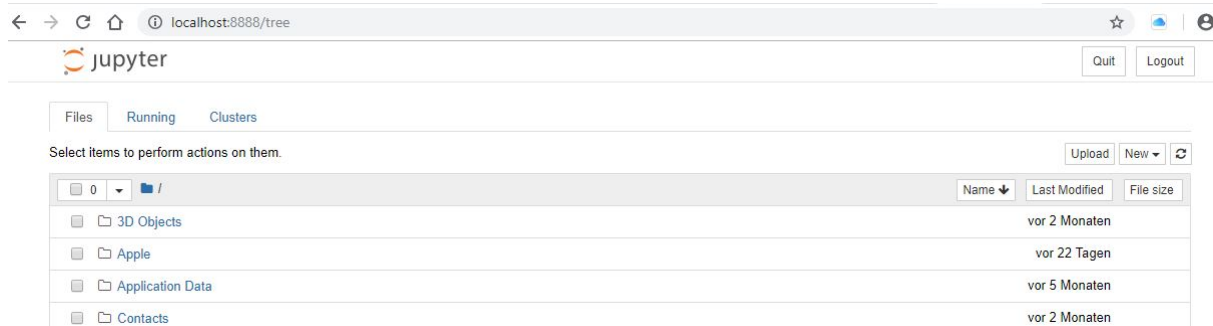


Abbildung 3: Jupyter - erstes Window

näher eingehen. Falls es Sie interessiert, empfehlen wir Ihnen, dieses Thema im Internet zu recherchieren.

Um ein neues Notebook anzulegen, klicken Sie auf New und wählen die Option Python 3 (siehe Abbildung 4). Sie sollten auf dem Bildschirm etwas sehen wie das in Abbildung 5. Machen Sie dies zum ersten Mal, klicken Sie probenhalber einmal auf die leere »Codezelle« und geben eine Zeile Python-Code ein. Zum Ausführen des Codes drücken Sie dann Shift-Enter.

Wenn Sie das Notebook speichern (siehe Save and Checkpoint im File-Menü des Notebooks), wird eine Datei mit der Dateierweiterung .ipynb angelegt. Das ist ein eigenständiges Dateiformat, das den gesamten momentanen Inhalt des Notebooks enthält (einschließlich aller evaluierten Codeausgaben). Dieser kann dann von anderen Jupyter-Benutzern geladen und bearbeitet werden. Um ein vorhandenes Notebook zu laden, kopieren Sie die Datei in das Verzeichnis, aus dem heraus Sie auch den Notebook-Prozess gestartet haben (oder ein Unterverzeichnis davon), und doppelklicken auf den Namen in der Startseite. Ein beispielhaftes Notebook sieht wie in Abbildung 6 aus.



Abbildung 4: Jupyter Notebook anlegen

Aufgabe 1 (Fortsetzung)

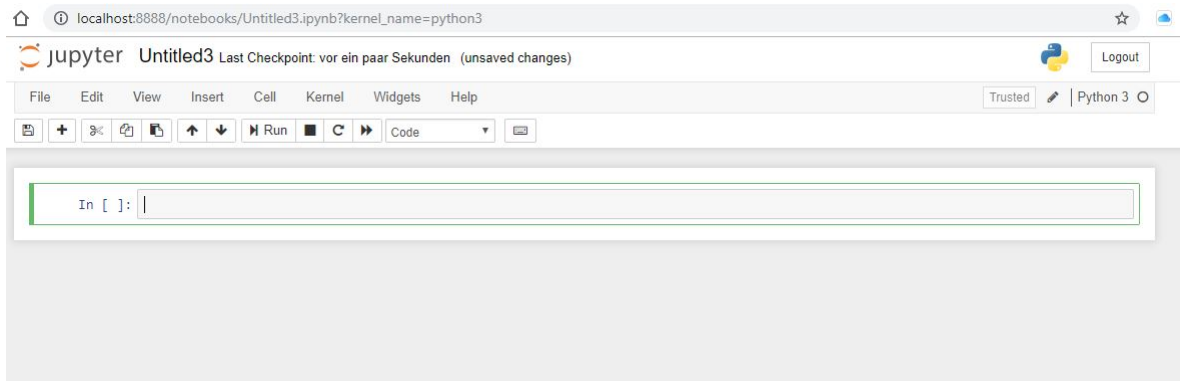


Abbildung 5: Jupyter neues leeres Notebook

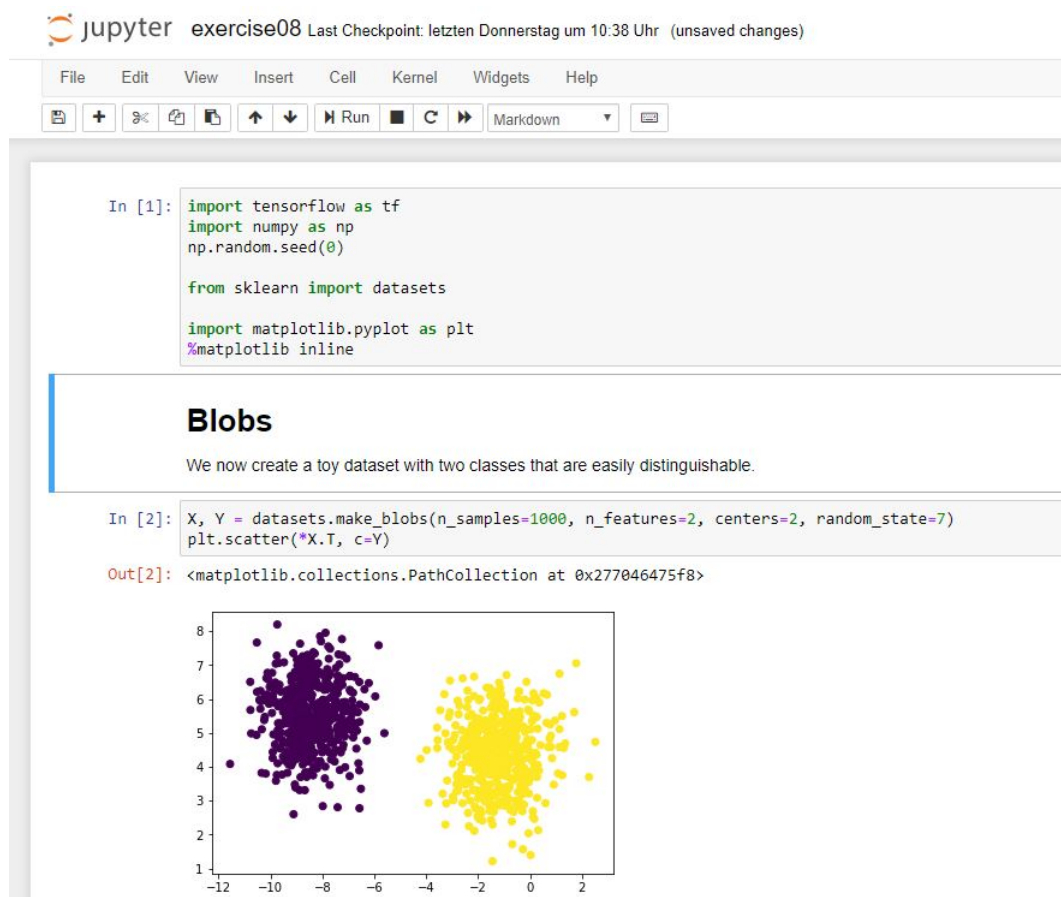


Abbildung 6: Jupyter Beispiel-Notebook

___ / 10 P.

Kleinklasse
Modulkürzel: w.BA.XX.2Stat-WIN.XX

Kurs: Wirtschaftsinformatik Fach: Statistik

Aufgabe 1 (Fortsetzung)

Aufgabe 2 (Modulendprüfung HS 10/11)

10 Punkte

Die Wartezeiten der Anrufenden in einem Callcenter wurden erfasst. Die Auswertung eines Tages ergab folgende Werte:

Wartezeit (in Min.)	0 < 2	2 < 4	4 < 6	6 < 8	8 < 10	10 < 12
# Anrufe	39	28	14	11	6	2
F(x _i)	0.39	0.67	0.81	0.92	0.98	1

Abbildung 7: Datenquelle: Wartezeiten im Callcenter

- (a) Ergänzen Sie obige Tabelle mit den kumulativen relativen Häufigkeiten F(x_i)!
 (b) In welche Klasse fällt der Median? (Kurze Begründung) **Bei 2<4 da Klasse 0<2 und 2<4 zusammen bereits 67% ergeben**
 (c) Wie hoch ist die mittlere Wartezeit in der Stichprobe (arithmetisches Mittel)?
 (d) Berechnen Sie die Stichprobenvarianz!
 (e) Bestimmen Sie die Modusklasse!

___ / 10 P.

c) mittlere Wartezeit (arithmetisches Mittel)

Man berechnet die Durchschnittswerte pro Wartezeit
 sprich
 1,3,5,7,9,11
 und rechnet mal #Anrufe

$$\frac{1 \times 39 + 3 \times 28 + 5 \times 14 + 7 \times 11 + 9 \times 6 + 11 \times 2}{100}$$

$$= 3.46$$

d) Berechnung Stichprobenvarianz:

e) Modusklasse bestimmen:

Aufgabe 3 (Tabellen und Diagramme)

10 Punkte

(Quelle 1: Aufgabe 1.4) Aus einer persönlichen Datensammlung eines Dozierenden stammt die folgende Häufigkeitsverteilung, die die Anzahl SMS pro Nachricht darstellt. (Eine SMS kann höchstens 160 Zeichen umfassen. Eine Nachricht, die mehr Zeichen enthält, wird in zwei oder mehr SMS aufgesplittet.) Die Tabelle gibt die Folgeglieder p_0 bis p_{10} an. Stellen Sie die Folge p_n graphisch dar.

Anzahl SMS pro Nachricht	Absolute Häufigkeiten	Relative Häufigkeiten	Kumulierte relative Häufigkeiten
1	349	0.68	0.68
2	114	0.22	0.9
3	36	0.07	0.97
4	10	0.02	0.99
5	3	0.006	0.996
6	2	0.004	1

514

Abbildung 8: Datenquelle: Eigene Erhebungen

- (a) Um welchen Datentyp handelt es sich bei diesem Merkmal? **diskrete metrische Daten**
- (b) Wie viele Nachrichten gab es im Beobachtungszeitraum? **514**
- (c) Wie viele SMS gab es im Beobachtungszeitraum? **752**
- (d) Ergänzen Sie die Tabelle mit den relativen und den kumulierten relativen Häufigkeiten!
- (e) Stellen Sie die relativen Häufigkeiten mit einem Stabdiagramm (Balkendiagramm) dar!
- (f) Stellen Sie die kumulierte relative Häufigkeitsverteilung graphisch dar!

___ / 10 P.

.

Aufgabe 4 (Tabellen und Diagramme)

6 Punkte

(Quelle 1: Aufgabe 1.6) Betrachten Sie folgende Häufigkeitsverteilung der monatlichen Nettolöhne von vollzeitbeschäftigten Arbeitnehmenden in der Schweiz nach Lohnhöhenklassen aus dem Jahre 2010 (privater und öffentlicher Sektor zusammen):

Lohnhöhenklassen (in CHF)	Total	Frauen	Männer
0 bis 3'000	2.3	5.2	1.1
3'001 bis 4'000	12.6	24.6	7.9
4'001 bis 5'000	23.7	26.5	22.6
5'001 bis 6'000	20.9	16.8	22.4
6'001 bis 7'000	12.9	10.2	13.9
7'001 bis 8'000	8.3	6.0	9.1
8'001 und mehr	19.3	10.7	23.0
Total	100.0	100.0	100.0

Abbildung 9: Datenquelle: Bundesamt für Statistik, Lohnstrukturerhebung 2010

- (a) Um was für einen Datentyp handelt es sich beim Merkmal «Nettolohn»? stetige metrische Daten
- (b) Stellen Sie die Lohnverteilung der Frauen mit einem Histogramm dar. Treffen Sie hierfür die Annahme, dass die letzte Klasse von CHF 8'001 bis CHF 12'000 reicht!

___ / 6 P.

Aufgabe 5 (Zusatz-Aufgabe)

6 Punkte

(Quelle 1: Aufgabe 2.3) An einer Hochschule wird ein Kurs nach der ECTS-Skala bewertet, die von A (relativ beste Ergebnisse) bis F (relativ schlechteste Ergebnisse) reicht. Die Noten der 15 Kursteilnehmer sind: A, B, C, B, C, B, A, F, B, C, D, D, C, B, B

- (a) Welche Skalierung weist das Merkmal «Note» auf? **ordinalskala**
- (b) Bestimmen Sie den Modus der Verteilung! **B**
- (c) Bestimmen Sie den Median der Verteilung!

___ / 6 P.