

Kursplan – Data Scientist, Utbildnings nr YH01458-2024-1 / YH01076-2024-1

Kurs: Pythonprogrammering och Statistisk Dataanalys (samt yrkesrollen)	Poäng: 45 yhp	Utgåva: 1
Framtagen av UL granskad av RUC: Monika Sandhaven/Molly Tagesson/Stefan Olsson	Språk: Svenska/Engelska	Datum: 2024-05-23
Förkunskaper: Inga förkunskaper	Granskad/Fastställd av: Ledningsgruppen	Sidan 1 (2)

Kursens huvudsakliga innehåll

Syftet med kursen är att den studerande ska få de kunskaper, färdigheter och kompetenser inom Python-programmering som krävs för att kunna arbeta med kvalificerad dataanalys. Den studerande kommer att få bekanta sig med bibliotek så som Numpy, Pandas och Matplotlib. De studerande kommer även att arbeta med grunderna i linjär algebra och statistik innefattande sannolikhetsteori och inferensteori samt implementera detta i Python. De studerande kommer även att få en introduktion i agila arbetssätt, github och datacamp. Utöver detta kommer den studerande lära sig att nyttja AI (Artificiell Intelligens) såsom generativ AI i arbetsprocessen.

Kursen omfattar följande moment:

- Användning av Pythons inbyggda typer
- Organisera och strukturera kod enligt kodstandard
- Objektorienterad programmering såsom nyttjandet av klasser
- Versionshantering och GitHub
- Grunderna i sannolikhetsteori och statistisk inferensteori
- Grunderna i linjär algebra
- Paket för dataanalys såsom, numpy, pandas, matplotlib
- Utföra beräkningar och visualisera data
- Agila arbetssätt
- Datacamp
- Användning av AI verktyg
- Yrkesrollen

Kursens mål/läranderesultat

Kunskaper:

1. Redogöra för pythons inbyggda datatyper, kontrollstrukturer, funktioner och grundläggande objektorientering samt de i branschen vanligt förekommande biblioteken
2. Förklara grunderna i sannolikhetsteori såsom att förstå och exemplifiera: diskreta och kontinuerliga fördelningar som t.ex. binomialfördelningen och normalfördelningen, väntevärde, varians och kovarians.
3. Redogöra för grunderna i statistisk inferensteori såsom konfidentintervall och hypotesprövning
4. Redogöra för grunderna i linjär algebra såsom matriser och vektorer

Färdigheter:

5. Tillämpa grundläggande pythonprogrammering för att skriva program som utför beräkningar och visualisering data.
6. Använda sannolikhets teori och statistisk inferens teori som verktyg för att analysera och dra slutsatser i olika beslutssituationer
7. Använda de i branschen vanliga biblioteken för dataanalys och maskininlärning på ett effektivt sätt

Former för undervisning

Kursen kommer att genomföras med blended learning med inspelningar, aktiva lektioner samt grupparbete.

Former för kunskapskontroll

Examination kommer att ske genom:

- 1 inlämningsuppgift som görs i grupp (IG/G)
- 2 inlämningsuppgifter (IG/G/VG).

Betygsskala

Följande betygsskala tillämpas:

VG = Väl Godkänd, G = Godkänd, IG = Icke Godkänd

Principer för betygssättning

Läranderesultat	Inlämningsuppgift 1 (G)	Inlämningsuppgift 2 (G/VG)	Inlämningsuppgift 3 (G/VG)
1	X		
2		X	
3			X
4	X		
5	X		
6		X	X
7	X		

För betyget Godkänd ska den studerande

- Ha nått samtliga läranderesultat för kursen

För betyget Väl Godkänd ska den studerande:

- Uppnått kraven för betyget Godkänd
- Använda sannolikhets teori och statistisk inferens teori för att lösa uppgifter med hög säkerhet och väl underbyggda resonemang

Icke Godkänt ges till studerande som har fullföljt kursen men inte nått alla mål för kursen.

Utbildare

Utbildare: Antonio Prgomet och Linus Rundberg Streuli.

E-post: Omniway.

Tillgänglighet: Möjligheten att ställa frågor och diskutera med utbildarna sker på lektionstid och vid behov via mejl funktionaliteten på Omniway. Försök att nyttja lektionstiden framför mejl för kontakt.

Kursmaterial

Typ av material	Kommentar
Kursens GitHub sida: https://github.com/AntonioPrgomet/ds24_pystat_kurs	
Kursbok: Lär dig Python från grunden av Antonio Prgomet och Linus Rundberg Streuli. ISBN: 9789181110500 Boken kan köpas på t.ex. Bokus. https://www.bokus.com/bok/9789181110500/lar-dig-python-fran-grunden/	
Statistisk Dataanalys av Svante Körner och Lars Wahlgren. ISBN: 9789144108704 Boken kan köpas på t.ex. Bokus. https://www.bokus.com/bok/9789144108704/statistisk-dataanalys/	Videor till boken: https://www.youtube.com/playlist?list=PLgzaMbMPEHEwkC-XVv3gpPrOk7y2IHWLJ
Tabeller och formler för statistiska beräkningar av Svante Körner och Lars Wahlgren. ISBN: 9789144114545 Boken kan köpas på t.ex. Bokus. https://www.bokus.com/bok/9789144114545/tabeller-och-formler-for-statistiska-berakningar/	
Kursbok: Matematik för yrkeshögskolan av Antonio Prgomet och André Enggård. ISBN: 9789181110098 . Boken kan köpas på t.ex. Bokus. https://www.bokus.com/bok/9789181110098/matematik-for-yrkeshogskolan/ Videor till boken: https://www.youtube.com/playlist?list=PLgzaMbMPEHEyWiQIOVHnR3FZzoUBEMB1A	Vi kommer i denna kursen använda kapitel 10 från denna boken.
Kolla på följande 8 videor om studieteknik, de kommer hjälpa dig mycket under dina studier och även senare i arbetslivet när du skall lära dig nya saker. https://www.youtube.com/watch?v=gSbpRjxYq24&list=PLA09CC1B5671827AD	

Kunskapskontroll – Regler

Resultat och betyg registreras senast inom 10 arbetsdagar från deadline.

Missar du ordinarie examination kontakta din Utbildningsledare och informera varför.

Missar man ordinarie deadline för kunskapskontrollen eller inte klarar den så är andra chansen 3 veckor efter ordinarie deadline. Du kan lämna in när du vill under denna period.

Missar man deadline för andra chansen eller inte klarar den så är tredje chansen 7 veckor efter ordinarie deadline. Du kan lämna in när du vill under denna period.

Har du missat samtliga examinationstillfällen så måste du kontakta din utbildningsledare snarast.

Upplägg på Föreläsningar / Lektioner

Lektionerna kommer fokusera på genomgångar och lösningar av uppgifter. Ert behov styr vad vi går igenom så skriv ned frågor som uppstår när ni studerar inför lektionerna så tar vi dem på lektionstid. Viktigt att man *inför* lektionerna arbetat med det material som förväntas enligt planeringen.

Schema:

Följande schema gäller för kursen mellan v. 40 – v.44.

	Förmiddag: 08.15 – 12.00.	Eftermiddag: 13.15 – 17.00
Måndag	Egenstudier för båda orter	Egenstudier för båda orter
Tisdag	Helsingborg / Malmö	Göteborg
Onsdag	Egenstudier för båda orter	Egenstudier för båda orter
Torsdag	Göteborg	Helsingborg / Malmö
Fredag	Egenstudier för båda orter	Egenstudier för båda orter

Följande schema gäller för kursen mellan v.45 – v48.

	Förmiddag: 08.15 – 12.00.	Eftermiddag: 13.15 – 17.00
Måndag	Egenstudier för båda orter	Egenstudier för båda orter
Tisdag	Helsingborg / Malmö	Göteborg
Onsdag	Helsingborg / Malmö	Göteborg
Torsdag	Egenstudier för båda orter	Egenstudier för båda orter
Fredag	Egenstudier för båda orter	Egenstudier för båda orter

Veckoplanering – Vad skall jag göra varje arbetsdag?

I detta avsnitt så framgår i detalj vad som skall göras varje dag. Viktigt att du följer schemat.

	Kursvecka 1 (v.40)
Mån	<p>1. Om du inte redan har kollat på följande video så kolla på den nu för att få en överblick av Python: https://www.youtube.com/watch?v=M7bnYJyCx0Q&list=PLgzaMbMPEHEy8nNzFoBxU-71bCwR_FPYv&index=1&t=1933s</p> <p>2. Läs dokumentet "mappstruktur".</p> <p>3. Se följande vdeo, fram till 06:55, om terminalen som är ett användargränssnitt för att interagera med datorn: https://www.youtube.com/watch?v=QcXBW-6FtrU&list=PLgzaMbMPEHEwe7nutbXE_oACkkXvAAaJ9&index=5&t=404s</p> <p>4. Se följande video, fram till 34:21, för att lära dig Git och GitHub. Detta används bokstavligen inom hela IT branschen. https://www.youtube.com/watch?v=9ybFhiYHKDg&list=PLgzaMbMPEHEwe7nutbXE_oACkkXvAAaJ9&index=8&t=2726s</p> <p>5. Läs kapitel 1 i kursboken "Lär dig Python från grunden" och gör tillhörande uppgifter.</p>
Tis	<p>Läs kapitel 2 i "Lär dig Python från grunden" och gör tillhörande uppgifter.</p> <p>Lektion via Teams kl: 13.15 - 17.00.</p>
Ons	<p>Läs kapitel 3 och kapitel 4 i "Lär dig Python från grunden" och gör tillhörande uppgifter.</p>
Tor	<p>Lektion via Teams kl: 08.15 - 12.00.</p>
Fre	<p>Läs kapitel 5 i "Lär dig Python från grunden" och gör tillhörande uppgifter.</p> <p>Skumma igenom kapitel 6 i "Lär dig Python från grunden". Det är ett mer utmanande kapitel.</p>

	Kursvecka 2 (v.41)
Mån	Läs kapitel 6 i "Lär dig Python från grunden" och gör tillhörande övningsuppgifter.
Tis	Läs kapitel 7 i "Lär dig Python från grunden" och gör tillhörande uppgifter. Lektion via Teams kl: 13.15 - 17.00.
Ons	Läs kapitel 10 "Vektor- och matrisalgebra" i boken "Matematik för yrkeshögskolan" och gör samtliga uppgifter.
Tor	Lektion via Teams kl: 08.15 - 12.00.
Fre	<ul style="list-style-type: none"> - Kolla igenom videon kopplat till NumPy: https://www.youtube.com/watch?v=M7bnYJyCx0Q [1:14:59 till 1:17:30]. -Läs länkarna nedan och experimentera med koden: - https://numpy.org/doc/stable/user/quickstart.html - https://numpy.org/doc/stable/user/basics.broadcasting.html - Arbeta med programmeringsuppgifterna i "kunskapskontroll_1_1_NumPy".

	Kursvecka 3 (v.42)
Mån	<p>Kolla igenom videon kopplat till Matplotlib: https://www.youtube.com/watch?v=M7bnYJyCx0Q [1:17:30 till 1:23:31].</p> <p>Läs länkarna och experimentera med koden: - https://matplotlib.org/stable/tutorials/introductory/usage.html#sphx-glr-tutorials-introductory-usage-py - https://matplotlib.org/stable/tutorials/introductory/pyplot.html#sphx-glr-tutorials-introductory-pyplot-py - Arbeta med programmeringsuppgifterna i "kunskapskontroll_1_2_matplotlib".</p> <p>Kolla på följande video om Pandas: https://www.youtube.com/watch?v=ZoNFPQUUsyk&t=28s och arbeta med tillhörande kod från videon.</p> <p>Läs dokumentationen "10 minutes to Pandas" och experimentera med koden: https://pandas.pydata.org/docs/user_guide/10min.html</p> <p>- Arbeta med programmeringsuppgifterna i "kunskapskontroll_1_3_Pandas".</p>
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	Arbeta med kunskapskontroll 1.
Tor	<p>Lektion via Teams kl: 08.15 - 12.00.</p> <p>Deadline kunskapskontroll 1 kl: 23.59 som består av de tre delarna NumPy, Matplotlib och Pandas som ni jobbat med tidigare.</p>
Fre	<p>Videorna hänvisar till följande spellista: https://www.youtube.com/playlist?list=PLgzaMbMPEHEwkc-XVv3gpPrOk7y2IHWLJ</p> <p>- Kolla video 1: Introduktion - Sannolikhetsteori & Statistisk Inferens - Kolla video 2: Introduktion - Sannolikhetsteori</p> <p>Läs kapitel 1 i kursboken. - Avsnitt 1.4 "Odds" är mindre viktigt, läs det översiktligt.</p> <p>Arbeta med uppgifterna: - 101-104 - 106-107 - 109-114 - 115-121</p>

	Kursvecka 4 (v.43)
Mån	<p>Kolla video 3: Sannolikhetsbegreppet.</p> <p>Läs kapitel 2 i kursboken.</p> <p>Arbeta med uppgifterna:</p> <ul style="list-style-type: none"> - 201-203 - 204-207 - 208-213 - 214-217 - 218-221 - 222-227
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	Repetition.
Tor	Lektion via Teams kl: 08.15 - 12.00.
Fre	<p>Kolla video 4 "Diskret slumpvariabel".</p> <p>Läs kapitel 3 i kursboken.</p> <p>Arbeta med uppgifterna:</p> <ul style="list-style-type: none"> - 301 - 304, 305 - 306 - 307-313 - 314-316 - 317 - 318-320

	Kursvecka 5 (v.44)
Mån	Repetition.
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	<p>Kolla video 5: "Tvådimensionell Slumpvariabel". Läs kapitel 4 i kursboken.</p> <p>Arbeta med uppgifterna: - 401-404, 407-412, 413 (I uppgift 413 härleder ni väntevärdet och variansen för Binomialfördelningen som är en summa av Tvåpunktsfördelade/Bernoulli slumpvariabler), 414-416.</p> <p>Kolla video 10 "Kod Demonstration - Statistik" fram till 04:32. Arbeta med koden "sannolikheteasteori_python"</p>
Tor	Lektion via Teams kl: 08.15 - 12.00.
Fre	<p>Kolla video 6: "Normalfördelningen". Läs kapitel 5 i kursboken.</p> <p>Arbeta med uppgifterna: - 501-508 - 509-512 - 513-517</p>

	Kursvecka 6 (v.45)
Mån	Repetition, särskilt kapitel 5 i kursboken som är väldigt viktigt.
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	<p>Kolla video 7 "Slumpmässigt urval och punktskattning". Läs kapitel 6 i kursboken.</p> <p>Arbeta med uppgifterna: - 601-609</p> <p>Lektion via Teams kl: 13.15 - 17.00.</p>
Tor	<p>Kolla video 8 "Konfidensintervall" och läs kapitel 7 i kursboken.</p> <p>Arbeta med uppgifterna: - 701, 702 - 703, - 704, 705 - 706 - 710</p>
Fre	Repetition.

	Kursvecka 7 (v.46)
Mån	Repetition.
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	Lektion via Teams kl: 13.15 - 17.00.
Tor	<p>Kolla video 9: Hypotesprövning Läs kapitel 8 kursboken.</p> <p>Arbeta med uppgifterna:</p> <ul style="list-style-type: none"> - 801-804 - 805 - 806 - 807 - 808, 809 - 810, 811 - 812 - 816
Fre	<p>Kolla video 9: Hypotesprövning Läs kapitel 8 kursboken.</p> <p>Arbeta med uppgifterna:</p> <ul style="list-style-type: none"> - 801-804 - 805 - 806 - 807 - 808, 809 - 810, 811 - 812 - 816

	Kursvecka 8 (v.47)
Mån	Repetition.
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	<p>- Kolla video 10 "Kod Demonstration - Statistik". - Arbeta med koden "inferens_python"</p> <p>Lektion via Teams kl: 13.15 - 17.00.</p>
Tor	Studera inför kunskapskontrollen i Sannolikheteori och Statistisk Inferens.
Fre	Studera inför kunskapskontrollen i Sannolikheteori och Statistisk Inferens.

	Kursvecka 9 (v.48)
Mån	Studera inför kunskapskontrollen i Sannolikhetsteori och Statistisk Inferens.
Tis	Lektion via Teams kl: 13.15 - 17.00.
Ons	Lektion via Teams kl: 13.15 - 17.00.
Tor	Studera inför kunskapskontrollen i Sannolikhetsteori och Statistisk Inferens.
Fre	Deadline kunskapskontroll 2 kl: 23.59.