

# Analys av Platsannonser från Arbetsförmedlingen – en datadriven pipeline i Python



Ana B.

EC Utbildning

Kunskapskontroll 2 Python fördjupad

202509

# Inledning

Arbetsmarknaden är en av de mest dynamiska delarna av samhället. Vilka yrken som efterfrågas, i vilka regioner, och hur behoven förändras över tid är frågor som är centrala för både arbetssökande, arbetsgivare, utbildningsanordnare och beslutsfattare. Att förstå dessa rörelser kan bidra till bättre matchning mellan arbetskraft och arbetsgivare, mer träffsäkra utbildningsinsatser och en övergripande bättre fungerande arbetsmarknad.

Traditionellt har analyser av arbetsmarknaden ofta byggts på historiska data som samlas in, bearbetas och publiceras med viss fördröjning. Det innebär att det finns en risk att man fattar beslut på information som redan är föråldrad. Genom att utnyttja öppna API:er, som det som Arbetsförmedlingen tillhandahåller via JobTech, finns möjlighet att analysera arbetsmarknaden i realtid eller nära realtid.

## Syftet med projektet

Syftet med detta projekt var att undersöka hur man kan bygga en datadriven pipeline för att löpande samla in och analysera platsannonser från Arbetsförmedlingen. Poängen är att på ett enkelt och automatiserat sätt kunna skapa en aktuell bild av arbetsmarknaden och visa:

- Vilka yrkesgrupper som har flest annonser just nu.
- Hur efterfrågan skiljer sig mellan olika län.
- Vilka trender som går att se under en kortare period (t.ex. 7 dagar).
- En första indikation på hur utvecklingen kan tänkas se ut nästa vecka genom en prognos.

## Bakgrund till projektet

Det finns flera anledningar:

- *Nytta för arbetsmarknaden:* Ett sådant här verktyg kan vara värdefullt för att identifiera kompetensbrist, planera utbildningsinsatser och förstå regionala skillnader.
- *Datadriven insikt:* Det visar på kraften i att använda öppna data i kombination med relativt enkel teknik (Python, SQLite, Matplotlib).
- *Praktiskt lärande:* För mig som student och utvecklare har projektet varit ett sätt att praktiskt tillämpa ETL-processer (Extract–Transform–Load), databaslagring, dataanalys och visualisering.
- *Framtidsperspektiv:* Genom att lägga till en prognosmodell visar projektet att det finns potential att gå från ren beskrivande analys till prediktiva analyser.

## Vad man kan visa med detta arbete :

Resultaten från pipeline kan användas för att:

- Visa den totala utvecklingen av annonser över tid.
- Rangordna yrkesgrupper utifrån antal annonser.
- Identifiera var i landet olika yrken är mest efterfrågade.
- Upptäcka kortsiktiga trender (7 dagar).
- Ge en första enkel prognos över kommande veckas behov.

Projektet visar därmed inte bara nuläget, utan även en metod för att löpande följa arbetsmarknaden och kommunicera resultaten på ett tydligt och visuellt sätt.

## Projektets upplägg :

För att göra analysen så robust och användbar som möjligt byggdes ett datadrivet flöde (pipeline) i fyra steg:

1. **Insamling** – Platsannonser hämtas via Arbetsförmedlingens JobTech API. Systemet klarar även fallback till testdata om API:t inte är tillgängligt.
2. **Lagring** – Annonserna sparas i en SQLite-databas, både som rådata och i aggregerad form per dag, yrkesgrupp och län.
3. **Bearbetning** – Data normaliseras (datumformat, länsnamn, yrkesgrupper) och görs jämförbara över tid.
4. **Analys och visualisering** – Resultaten presenteras i diagram och tabeller som gör det lätt att upptäcka mönster.

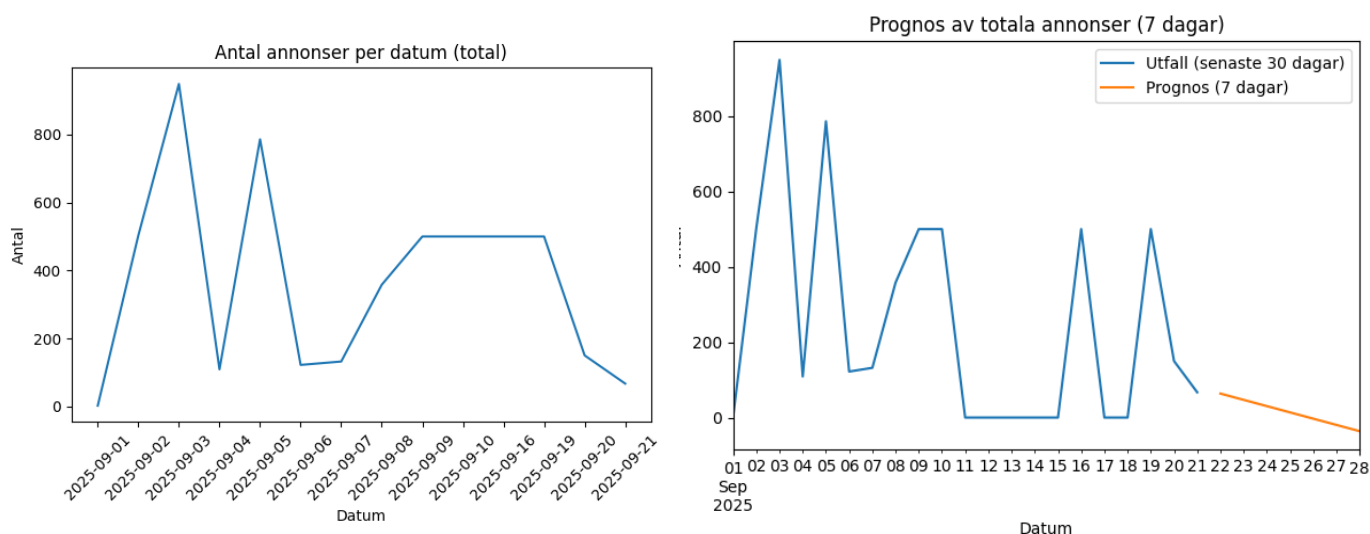
En extra komponent är en **mini-prognos** (7 dagar framåt) som bygger på statistiska modeller. Den gör det möjligt att inte bara analysera nuläget utan även blicka framåt.

## Resultat och insikter

### Totalt antal annonser

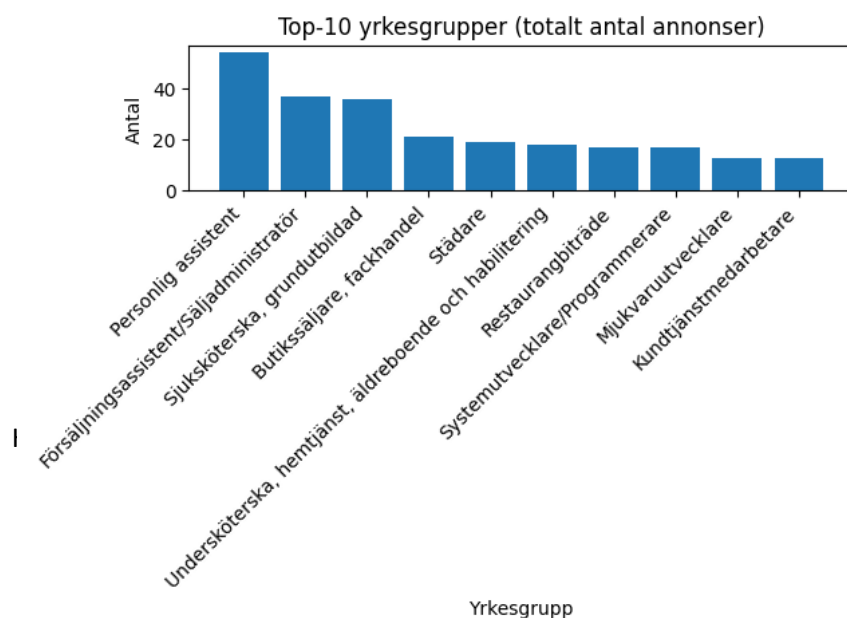
Diagrammet nedan visar utvecklingen av det totala antalet annonser samt en prognos för de kommande sju dagarna. Prognosen bygger på historiska data och ger en indikation på om arbetsmarknadsaktiviteten väntas öka eller minska.

Trots korta tidsserier kan man redan se hur annonserna fluktuerar, och prognosen antyder en stabilisering runt en viss nivå.



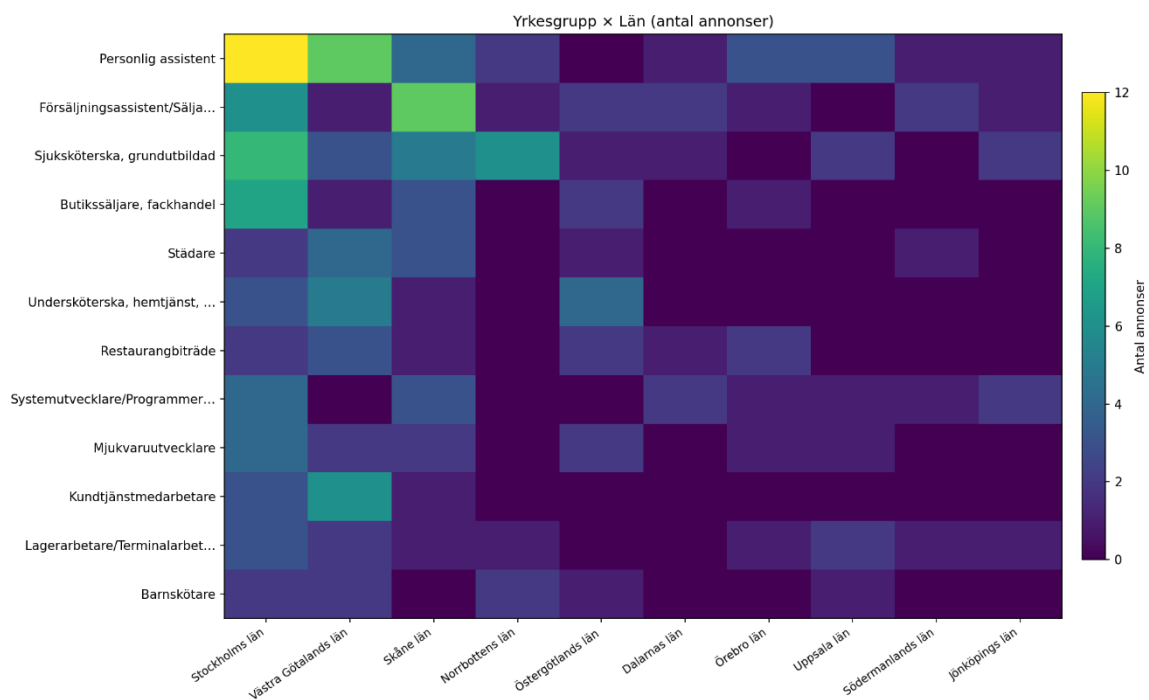
### Yrkesgrupper och län

För att förstå vilka yrken som efterfrågas mest användes en **Top-10-analys**. Diagrammet visar de tio yrkesgrupper som har flest annonser.



- **Personlig assistent** har den största volymen.
- **Försäljningsassistent/Säljadministratör** och **Sjuksköterska** är också starkt representerade.
- Yrkesgrupper inom vård, omsorg och IT dominerar.

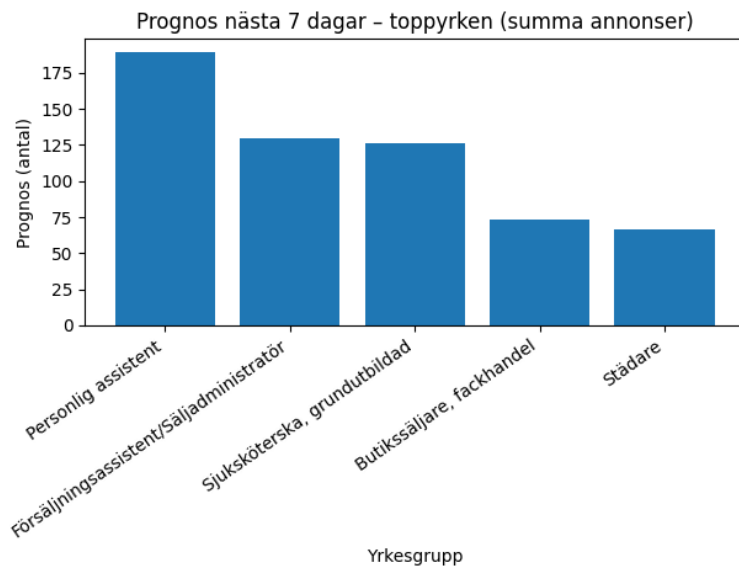
För att kombinera yrkesgrupper och regioner användes en värmekarta. Den visar i vilka län olika yrken är mest efterfrågade.



Stockholm och Västra Götaland sticker ut, särskilt inom IT-yrken, medan vårdyrken finns brett i flera län.

## Prognos för toppyrken

För att testa potentialen i prediktion togs en 7-dagars prognos fram för de mest aktiva yrkesgrupperna. Resultatet visar förväntat antal annonser nästa vecka.



Enligt prognosen kommer **Personlig assistent** fortsätta dominera, följt av **Försäljningsassistent/Säljadministratör** och **Sjuksköterska**. Detta speglar tydliga trender inom både omsorgssektorn och serviceyrken.

## Slutsatser

Projektet visar att det är möjligt att bygga en enkel men kraftfull pipeline för att analysera arbetsmarknaden i realtid:

- Arbetsförmedlingens öppna data kan användas för att identifiera de mest efterfrågade yrkena och regionerna.
- Visualiseringar gör trender och mönster lätta att upptäcka och kommunicera.
- En första prognosmodell ger en framåtblickande dimension, även om datan än så länge är begränsad.

## Möjliga nästa steg

- Automatisera daglig körning av pipeline.
- Bygga längre tidsserier för mer tillförlitliga prognoser.
- Utveckla ett interaktivt dashboard-gränssnitt

# 1 Självutvärdering

## 1. Utmaningar du haft under arbetet samt hur du hanterat dem.

Den största utmaningen har varit att förstå och tolka Arbetsförmedlingens API och att på ett konsekvent sätt hantera den data som hämtas. Initialt var det svårt att se hur datan var strukturerad och vilka fält som var mest relevanta för analys. För att lösa detta arbetade jag iterativt: jag började med sample-data, gick vidare till riktiga API-anrop och lade sedan successivt till lagring i SQLite, transformationer och visualiseringar. Genom att bryta ner arbetet i mindre steg kunde jag isolera problem och lösa dem en efter en, vilket gjorde processen mer hanterbar.

## 2. Vilket betyg du anser att du skall ha och varför.

Jag anser att arbetet motsvarar kraven för VG. Jag har byggt en komplett ETL-pipeline som automatiskt hämtar data, lagrar den i en databas och genomför analyser och visualiseringar. Projektet innehåller både daglig datahämtning, trender över tid, topplistor och en prediktiv analysdel som visar möjliga utvecklingar för kommande vecka. Arbetet visar att jag kan kombinera datainsamling, databaslagring, bearbetning och visualisering i en sammanhängande lösning.

Dessutom finns det dokumentation, rapport samt README för GitHub.

## 3. Något du vill lyfta fram till Antonio?

Projektet gav mig värdefull praktisk erfarenhet av att bygga en pipeline från grunden, där teori översattes till ett fungerande flöde. Jag vill särskilt lyfta fram hur viktigt det var att arbeta iterativt och testa varje delmoment innan jag gick vidare. På så sätt kunde jag tidigt få feedback på vad som fungerade och vad som behövde justeras. Jag ser också stor potential i att vidareutveckla arbetet, exempelvis genom att förbättra prognosmodellen eller lägga till fler interaktiva visualiseringar.

## Källförteckning

Arbetsförmedlingen / JobTech Dev – <https://jobtechdev.se/>