SAHRASH FATIMA

sahrashfatima8@gmail.com | 463-276-4476 |Open to relocation, IN
https://github.com/Sahrashfatima?tab=repositories
www.linkedin.com/in/sahrash-fatima

SUMMARY

- **Health Informatics Specialist** with over 4+ years of experience in clinical data analysis, public health informatics, and healthcare research.
- Proficient in leveraging **REDCap, OnCore,** and **OpenEMR** systems for secure data collection, clinical research management, and decision support implementation.
- Skilled in developing **FHIR-based ETL pipelines** using **Python, SQL,** and **JSON** to enable interoperability between EHR systems and external applications.
- Experienced in designing interactive dashboards using **Power BI and Tableau** to visualize healthcare trends, patient outcomes, and program KPIs.
- Adept at processing, cleaning, and analyzing large-scale datasets across research and hospital settings using Python, **R, Excel,** and MySQL.
- Familiar with **HL7, ICD-10, SNOMED CT**, and **LOINC** standards for clinical terminology and data integration.
- Knowledgeable in version control using **Git and GitHub**, and experienced in API testing, project documentation, and stakeholder communication.

SKILLS

- **Programming & Scripting:** Python, SQL, Java, Python Script, R , JSON, HL7, FHIR, LOINC, ICD- 10, SNOMED CT
- **Methodologies:** Agile, Waterfall, Wrike, BPMN, Gantt Chart
- **Data Visualization:** Power BI, Tableau, Microsoft Excel (Advanced, Pivot Tables, Macros, VBA, Data Modeling)
- **Libraries:** NumPy, Pandas, SciPy, PySpark, Matplotlib
- **Version Control Tools:** Git, GitHub, Postman, OAuth 2.0, Visual Studio, WordPress
- **Databases:** MySQL, SQL, PostgreSQL, Excel
- **ETL Tools:** Python based ETL, OpenEMR FHIR API, ETL Automation
- **Big Data & Machine Learning:** Logistic Regression, Data Modeling, Predictive Analytics, Time- series Analysis, NLP
- **Cloud Platforms:** AWS (S3, EC2, Redshift, IAM, Glue), SharePoint, REDCap Cloud, Teams
- **Data Analysis:** Statistical Analysis, Reporting, A/B Testing, Data Warehousing, Data Mining,
- **Data Warehousing**: Snokeflake Big Query, Data Modeling, Data Migration

PROFESSIONAL EXPERIENCE

Indiana University – Randolph County Caring Community Partnership (RCCCP) **May 2024 -**
Business Intelligence Analyst/ Data Analyst **Present USA**

- Configured **DHIS2** data elements and **public health workflows** based on stakeholder input, increasing data accuracy and **reporting** efficiency by 40%.
- Designed real-time Power BI dashboards to visualize **SDOH KPIs,** reducing manual data consolidation efforts by 60%.
- Automated reporting pipelines and cross-sector evaluation dashboards, enhancing program monitoring across 14+ counties.
- **Trained and mentored** 25+ interns on DHIS2 configuration and dashboard development, accelerating onboarding and improving project outcomes.
- Led testing and **API-based integration** of external systems with DHIS2, improving data interoperability and system adoption by 35%.
- Designed **Power BI dashboards**, aggregating data from 5+ sources (EHRs, DHIS2) to track KPIs like patient wait times for accepting into program.
- Automated **ETL workflows**, reducing manual data processing time from 12 hours to 4 hours weekly, generating Seaborn visualizations to analyze population health trends across 10K+ patient records.
- Applied advanced **SQL** optimizations, reducing query execution time from 60 seconds to 15 seconds via indexing and partitioning, accelerating analytics for real-time patient risk stratification.
- Leveraged **Excel (VLOOKUPs, Pivot Tables, automation)** for data reconciliation, enabling faster reporting and improving healthcare operational insights.
- Collaborated with cross-functional teams, including healthcare professionals and business stakeholders, to align data insights with operational goals, improving decision-making in patient care management.

IU School of Medicine – Riley Children's Hospital **April 2024 - Present**
Research Analyst **USA**

- Designed and deployed predictive models using **Python** and **SQL**, improving guest demand forecasting accuracy by 20% and optimizing pricing strategies.
- Conducted **REDCap and OnCore**-based data cleaning and merging for multi-site pediatric studies, reducing data processing time by 30%.
- Support ed grant submissions and internal planning. Contributed to IRB documentation, participant tracking, and qualitative interview protocols for mixed-method research.

- Engineered and enforced 6+ **Python-based scripts** for data extraction and transformation, enabling efficient project-specific analyses and improving data accessibility.
- Established **data governanc e** frameworks, ensuring compliance with internal policies and maintaining data quality standards across 5+ client projects for technology-driven solutions.
- Collaborated with the IT and clinical teamsot integrate data from 5+ training sites, streamlining the availability of critical patient and operational data.
- Facilitated **data mining** on 5,000+ patient reports, identifying padmission rates and treatment u, which informed predictive models for optimizing resource allocation in healthcare facilities training.
- Administered weekly data quality checks using **Azure Data Studio**, resolving inconsistencies and enhancing the reliability of healthcare reports.

Excelra – GOSTAR® Database Team **May 2023 - Dec 2023, India**
Scientific Researcher

- **Curated compound**, assay, and **SAR data** from **scientific publications** to support **AI-driven drug discovery pipelines**.
- **Validated** and structured chemical and biological data for **integration into GOSTAR®**, enabling **ML/QSAR model development**.
- Supported pharmacology data extraction across 10+ therapeutic areas, improving database usability for downstream analytics.

Rohini Multi-Specialty Hospital, India **Aug 2019 – May 2023**
Clinical Pharmacist/ Compliance Analyst

- Generated weekly medication usage and **outcome reports** from hospital records, supporting treatment planning and resource allocation.
- Conducted **pharmacovigilance tracking** and **ADR documentation**, contributing to longitudinal **drug safety evaluation**.
- Educated patients on medication adherence and reconciled prescriptions at discharge to prevent adverse drug interactions.

EDUCATION

Indiana University Indianapolis
Master in Health Informatics **GPA: 3.8**

Kakatiya University
Doctor of pharmacy **GPA: 9/10**
India

PROJECTS

Statistical Analysis for Parkinson's Disease Diagnosis Predictors
• Conducted advanced statistical analysis across six clinical and demographic domains (e.g., UPDRS, MoCA, tremor severity) and Built and validated a multiple logistic regression model (AUC = 0.8261), and visualized results using violin plots, heatmaps, and AUC curves to facilitate clinical interpretation of early-stage Parkinson's predictors.

Predictive Modeling of Hemodialysis Frequency in ESRD Patients
• Developed and compared machine learning classifiers (logistic regression, random forest, gradient boost) to predict dialysis frequency using features from CKD patient records including comorbidities, labs, and demographics.
• Performed extensive EDA and statistical validation (Shapiro-Wilk, Chi-Square, Mann-Whitney U), and visualized insights using heatmaps, line graphs, and pie charts to reveal key clinical patterns and feature gaps in Kaggle datasets.

Big Data Visualization: Obesity & Drug-Related Deaths
• Cleaned and segmented U.S. mortality datasets by drug type (opioid vs. non-opioid), race, and gender to support risk stratification and public health analysis. Designed interactive Power BI dashboards showing geographic and temporal trends in obesity and drug-related deaths, highlighting high-risk populations and regions for targeted interventions.

OpenEMR Clinical Decision Support System (CDSS) Implementation
• Developed two evidence-based Clinical Decision Rules (CDRs) within OpenEMR for Achalasia and FAS prevention, using ICD-10 coding, ClinicalKey references, and Excel-mapped care pathways. Integrated CDRs with logic-based triggers and alerts to support diagnosis and behavioral screening (e.g., alcohol misuse), enhancing clinical workflow and decision-making.

FHIR-Based ETL Pipeline for Healthcare Interoperability
• Built a secure ETL pipeline using Python and OpenEMR FHIR APIs to extract, transform, and load structured clinical data into a primary care EHR, enabling standards-based interoperability. Applied OAuth 2.0 authentication, SNOMED CT and LOINC terminologies, and BPMN use case modeling; visualized data coverage across U.S. states

CERTIFICATIONS

- CITI – Biomedical Research (Stage 1)
- HIPAA Compliance Training
- FERPA (Family Educational Rights & Privacy Act)
- Good Clinical Practice (GCP)
- REDCap
- Power BI
- OnCore