

Allegato “A”

Metodi Statistici e di Data Science per il decision – Making in Sanità: analisi dei dati sanitari (Data Analytics) RADAR corso base ARES Sardegna

Il presente allegato è composto
di n. 8 fogli, di n. 8 pagine

METODI STATISTICI E DI DATA SCIENCES PER IL DECISION- MAKING IN SANITÀ: ANALISI DEI DATI SANITARI (DATA ANALYTICS)

Corso RADAR Base

Destinatari	Personale della Struttura Complessa Programmazione e Controllo di Gestione e della Struttura Semplice Dipartimentale Innovation Management, Audit e Processi Organizzativi, con particolare riferimento a dirigenti, funzionari, collaboratori professionali, assistenti e analisti coinvolti nell'utilizzo operativo di dati sanitari, amministrativi, economici e gestionali.
Durata totale	20 ore - 4 giornate da 5 ore ciascuna
Modalità	• Giornate 3 e 4 in presenza presso la sede della Direzione Generale. • Giornate 1 e 2 erogabili anche a distanza o in modalità laboratoriale assistita.
Numero partecipanti	13 partecipanti già identificati, con possibilità di estensione ad altri professionisti di ulteriori Servizi aziendali, fino a un massimo di 20 partecipanti per corso.

Razionale

La capacità di utilizzare i dati in modo strutturato, tempestivo e affidabile è una competenza essenziale per il governo delle aziende sanitarie. Le aziende dispongono di molte basi informative - flussi amministrativi e clinico-amministrativi, produzione, costi, mobilità, liste d'attesa, contratti, magazzino e personale - ma spesso tali dati sono frammentati, non standardizzati e difficili da trasformare in informazioni utili per le decisioni. Il corso RADAR Base nasce per rafforzare le competenze operative interne sull'uso dei dati sanitari, con un taglio concreto e progressivo: prima la costruzione di basi dati affidabili in Excel, poi il calcolo di KPI e la reportistica tramite tabelle e grafici pivot, infine l'introduzione a Python e Codex come strumenti per automatizzare controlli, ridurre errori manuali e rendere replicabile il lavoro analitico.

Obiettivi formativi

Il corso Radar Base fornisce conoscenze teoriche e strumenti pratici per utilizzare i dati sanitari a supporto delle decisioni manageriali, economiche e organizzative. Al termine del percorso i partecipanti saranno in grado di:

- comprendere la logica dei principali dati amministrativi e clinico-amministrativi utilizzati in sanità;
- strutturare correttamente una base dati, evitando errori di formattazione, duplicazioni, codifiche incoerenti e problemi di tracciabilità;
- utilizzare Excel per pulizia, normalizzazione e controllo del dato, con funzioni logiche, di ricerca e di testo;
- applicare convalida dati, formattazione condizionale, tabelle strutturate, filtri e controlli di coerenza;
- costruire Tabelle Pivot, Grafici Pivot e report sintetici per KPI gestionali, economici e produttivi;
- leggere statistiche descrittive di base, scostamenti e anomalie rilevanti per la Direzione;
- impostare un ambiente di lavoro Python e utilizzare Codex o strumenti AI di coding per generare, correggere e documentare script semplici;
- automatizzare controlli di qualità e produrre output puliti e riutilizzabili per reportistica e analisi successive.

Modalità didattiche

Il corso adotta una metodologia mista, interattiva e laboratoriale. Ogni modulo alterna inquadramento metodologico, dimostrazione guidata e applicazione su dataset sanitari simulati, anonimizzati o generici.

- lezioni frontali brevi per acquisire metodo e lessico comune;
- laboratori guidati in Excel su dati sanitari e amministrativo-gestionali;
- esercitazioni su Tabelle Pivot, Grafici Pivot, formattazione condizionale e convalida dati;
- introduzione pratica a Python, Visual Studio Code/Jupyter e pandas;
- uso di Codex e AI generativa come tutor operativo per codice, controlli qualità e documentazione;
- discussione manageriale dei risultati e traduzione degli output analitici in insight decisionali.

Struttura del corso

Giornata 1

Dalla qualità del dato al governo manageriale dell'azienda sanitaria

Docenti: Fabrizio Schettini - Daniele Bellavia

Obiettivo della giornata

Fornire ai partecipanti un metodo per comprendere la logica dei dati amministrativi e gestionali in sanità, riconoscere le principali criticità di qualità del dato e impostare correttamente dataset utilizzabili per analisi, KPI e reporting direzionale.

- **La sanità data-driven: dal dato amministrativo alla decisione manageriale:** il dato come infrastruttura per governare produzione, costi, performance, mobilità, appropriatezza e processi aziendali.

- **I principali flussi informativi sanitari e gestionali:** SDO, File F, File G, specialistica, farmaceutica, mobilità sanitaria, contabilità analitica, centri di costo, dati di produzione e dati economico-organizzativi.
- **Qualità del dato e criticità ricorrenti:** completezza, coerenza, accuratezza, tempestività, tracciabilità e confrontabilità. Errori frequenti: codifiche incoerenti, duplicazioni, valori mancanti, formati non omogenei, campi descrittivi non standardizzati e problemi di raccordo tra fonti diverse.
- **Strutturazione operativa del database in Excel:** tabelle strutturate, corretta organizzazione di righe e colonne, intestazioni, formati numerici e date, filtri, ordinamenti, rimozione duplicati, separazione e normalizzazione di campi testuali, gestione di codici e chiavi di collegamento.
- **Strumenti Excel per prevenire errori e controllare la qualità del dato:** formattazione condizionale, convalida dati, elenchi a tendina, regole di controllo, evidenziazione automatica di valori mancanti o anomali, protezione delle celle critiche e impostazione di file più robusti e meno esposti a errori manuali.
- **Funzioni operative per il controllo qualità e la preparazione del dato:** utilizzo guidato di funzioni logiche, di ricerca e di normalizzazione, tra cui CERCA.VERT, CERCA.X, INDICE, CONFRONTA, PIÙ.SE, funzioni condizionali e funzioni di testo per controllare codifiche, riclassificare prestazioni, correggere descrizioni, individuare incongruenze e preparare basi dati utilizzabili per KPI e reporting.

Giornata 2

KPI sanitari, statistica descrittiva, analisi degli scostamenti e reportistica per la Direzione

Docenti: Fabrizio Schettini - Daniele Bellavia

Obiettivo della giornata

Consentire ai partecipanti di trasformare dati amministrativi, economici e produttivi in indicatori utili per il governo aziendale, introducendo anche gli strumenti essenziali di statistica descrittiva per leggere correttamente distribuzioni, variabilità, anomalie e scostamenti. La giornata si focalizza su KPI sanitari, mobilità sanitaria, centri di costo, budget, costi effettivi e reporting sintetico per la Direzione.

- **Dal dato elementare all'indicatore manageriale:** volumi, costi, tariffe, ricavi, mix produttivo, appropriatezza, saturazione, tempi, scostamenti e indicatori di performance.
- **Statistica descrittiva di base per leggere i dati sanitari:** media, mediana, deviazione standard, minimo, massimo, quartili, percentili e distribuzioni. Differenza tra media e mediana nei dati sanitari asimmetrici, ad esempio costi, durate di degenza, tempi di attesa, numero di accessi o consumo di risorse.
- **Tabelle Pivot per KPI e analisi multidimensionale:** costruzione di Tabelle Pivot per analizzare dati per reparto, centro di costo, periodo, prestazione, classe di priorità, regime, provenienza, mobilità attiva/passiva e linea di attività. Utilizzo di filtri, segmentazioni, campi calcolati e raggruppamenti.
- **Grafici Pivot e visualizzazione delle anomalie:** grafici Pivot, istogrammi, distribuzioni di frequenza, box plot, outlier e valori estremi. Come scegliere la visualizzazione più adatta per rendere leggibile un fenomeno alla Direzione.
- **Indicatori per il controllo di gestione sanitario e analisi degli scostamenti:** Lettura delle varianze e distinzione tra effetto volume, effetto mix, effetto costo ed effetto tariffa.

- **Laboratorio applicativo:** applicazione su dataset sanitario per costruire Tabelle Pivot, Grafici Pivot, KPI, statistiche descrittive e analisi degli scostamenti. I partecipanti lavorano su un caso guidato per produrre una breve sintesi direzionale con indicatori, anomalie, driver principali e possibili implicazioni manageriali.

Giornata 3

Introduzione a Python, AI e Codex per la qualità del dato e l'automazione dei controlli

Docenti: Sara Pozzi o Francesco Bertolotti; Fabrizio Schettini o Daniele Bellavia

Obiettivo della giornata

Introdurre i partecipanti all'utilizzo di Python e dell'AI generativa come strumenti di supporto per automatizzare controlli di qualità, individuare anomalie, gestire dataset sanitari di grandi dimensioni e ridurre attività manuali ripetitive. La giornata parte dalle basi, accompagnando i partecipanti nell'impostazione dell'ambiente di lavoro e nei primi passaggi operativi necessari per utilizzare Python in modo semplice, guidato e replicabile.

- **Perché introdurre Python:** grandi moli di dati, attività ripetitive, necessità di tracciabilità, riduzione dell'errore manuale, maggiore replicabilità delle analisi e superamento dei limiti degli strumenti tradizionali.
- **Impostare l'ambiente di lavoro partendo da zero:** installazione e orientamento a Python, Visual Studio Code, Jupyter Notebook o ambienti equivalenti. Struttura delle cartelle di lavoro, gestione dei file, apertura di un progetto, esecuzione dei primi comandi e salvataggio degli script.
- **Introduzione operativa a Python per non programmatori:** logica di base del linguaggio, variabili, librerie, importazione dei dati, lettura di file tabellari, prime esplorazioni del dataset e introduzione alla libreria pandas per la manipolazione dei dati.
- **Controlli automatici di qualità del dato:** identificazione di duplicati, valori mancanti, codici anomali, date incoerenti, formati non standardizzati, campi fuori range e possibili errori di codifica nei flussi sanitari.
- **AI generativa e Codex come assistenti operativi:** come formulare prompt efficaci per generare codice, correggere errori, spiegare passaggi complessi, documentare le trasformazioni e costruire checklist di controllo qualità. Verranno presentati esempi di utilizzo di strumenti AI per il coding, inclusi ambienti come Codex ove disponibili.
- **Laboratorio guidato**

Costruzione passo-passo di uno script Python per controllare e validare un flusso sanitario simulato. Il laboratorio prevede: importazione del dataset, controllo della struttura, individuazione di duplicati e valori mancanti, verifica di codici e date, produzione di un breve report di qualità del dato ed esportazione di un file pulito. L'attività sarà svolta con il supporto dell'AI generativa e, ove disponibile, di Codex come assistente operativo per generare, correggere e documentare il codice.

Giornata 4

Data Cleaning e consolidamento di fonti eterogenee con Python

Docenti: Sara Pozzi o Francesco Bertolotti; Fabrizio Schettini o Daniele Bellavia

Obiettivo della giornata

Superare i limiti degli strumenti tradizionali introducendo Python come supporto operativo per pulire, integrare e consolidare basi dati sanitarie provenienti da fonti diverse. La giornata si concentra sul collegamento tra dati di contabilità analitica e dati di produzione clinica, con l'obiettivo di costruire un dataset unico, controllato e pronto per alimentare report gestionali, tabelle di sintesi e successive dashboard direzionali.

- **Workflow replicabile di data cleaning:** impostazione di cartelle, file di input, script, file di output e controlli minimi per rendere tracciabile il processo di elaborazione.
- **Pulizia e standardizzazione con Python:** normalizzazione di codici, campi testuali, date, valori mancanti e duplicati, con esportazione di dataset puliti.
- **Introduzione al consolidamento di fonti:** unione semplice di archivi tramite chiavi comuni, gestione delle corrispondenze mancanti e controllo delle righe non abbinata.
- **Output per reportistica:** generazione di file puliti, tabelle riepilogative e report sintetici con indicatori principali e controlli di qualità.
- **Uso di Codex per migliorare lo script:** revisione del codice, spiegazione degli errori, documentazione dei passaggi, suggerimenti di refactoring e checklist finale di validazione
- **Laboratorio guidato:** sviluppo guidato di uno script che pulisce due o tre file sanitari/amministrativo-gestionali, li consolida tramite chiavi comuni, produce un file finale pulito e genera un breve report con controlli di qualità, record non abbinati e principali indicatori di sintesi.

Propedeuticità e valutazione finale

I due corsi sono progettati come percorsi propedeutici: RADAR Base fornisce le competenze operative preliminari su qualità del dato, Excel, KPI, Python e Codex; RADAR Advanced ne costituisce la prosecuzione naturale, con focus su R, Power BI, DAX, AI nativa e cruscotti decisionali. Pur non essendo previsto accreditamento ECM, al termine di ciascun corso sarà effettuata una valutazione finale dell'apprendimento, tramite test o prova applicativa coerente con gli obiettivi formativi.

Budget progettuale

Il Corso RADAR Base prevede un budget progettuale comprensivo dei costi vivi di erogazione, progettazione didattica, direzione scientifica, coordinamento, predisposizione e correzione delle esercitazioni, logistica dei docenti verso la sede del corso e kit di benvenuto. Tali costi sono riferiti alla realizzazione di un percorso formativo di 20 ore complessive, articolato in 4 giornate da 5 ore ciascuna, di cui 2 giornate in presenza presso la sede della Direzione Generale e 2 giornate erogabili a distanza o in modalità laboratoriale assistita, con taglio operativo-laboratoriale.

La voce logistica è stimata esclusivamente sulla base delle 2 giornate in presenza previste presso la sede della Direzione Generale.

Di seguito viene riportato il budget complessivo previsto per l'erogazione del corso.

VOCE DI BUDGET	ATTIVITÀ COMPRESSE	IMPORTO
LOGISTICA	Gestione della logistica dei docenti per le giornate in presenza presso la sede della Direzione Generale, inclusi spostamenti e pernottamenti.	€ 1.256,00
PROGETTAZIONE CORSO	Progettazione didattica e stesura del programma di dettaglio, con definizione della sequenza dei moduli, degli obiettivi formativi, dei contenuti e delle modalità laboratoriali.	€ 560,00
DIREZIONE SCIENTIFICA	Direzione scientifica e controllo dei contenuti e delle metodologie proposte, al fine di garantire coerenza tra fabbisogni formativi, obiettivi del corso, casi applicativi e strumenti didattici utilizzati.	€ 400,00
PREDISPOSIZIONE E CORREZIONE ESERCITAZIONI	Predisposizione di esercitazioni, casi applicativi, dataset di lavoro e materiali per lavori individuali o di gruppo; correzione degli elaborati e supporto alla valutazione degli apprendimenti.	€ 800,00
COORDINAMENTO DIDATTICO	Coordinamento didattico dei docenti coinvolti, definizione del materiale didattico, allineamento degli interventi, predisposizione del test iniziale e valutazione finale delle conoscenze acquisite.	€ 800,00
DOCENZE	Attività di docenza per 20 ore complessive di lezione, articolate in 4 giornate da 5 ore ciascuna, secondo la struttura didattica prevista dal programma.	€ 2.343,75
KIT DI BENVENUTO	Predisposizione di kit per i partecipanti.	€ 78,00
ONERI UNIVERSITARI	Utilizzo della piattaforma didattica online, gestione dei partecipanti, storage del materiale didattico, supporto organizzativo, gestione amministrativa e attività di back office collegate all'erogazione del corso.	€ 2.935,41
TOTALE COMPLESSIVO		€ 9.173,16

Progetto Formativo

Metodi Statistici e di Data Science per il decision – Making in Sanità: analisi dei dati sanitari (Data Analytics) – RADAR corso base – ARES Sardegna

TABELLA RIEPILOGATIVA COSTI PREVISTI

VOCE DI SPESA	TOTALE
Compenso docenze, logistica, progettazione corso, direzione scientifica, predisposizione e correzione esercitazioni, coordinamento didattico, kit discenti e oneri universitari	€ 9.173,16
TOTALE	€ 9.173,16

Costo medio per partecipante € 458,66
Costo medio x ora x partecipante € 22,93