

Spese mediche

Descrizione dataset

Descrizione del Dataset				
	age	bmi	children	charges
count	1337.00	1337.00	1337.00	1337.00
mean	39.22	30.66	1.10	13279.12
std	14.04	6.10	1.21	12110.36
min	18.00	15.96	0.00	1121.87
25%	27.00	26.29	0.00	4746.34
50%	39.00	30.40	1.00	9386.16
75%	51.00	34.70	2.00	16657.72
max	64.00	53.13	5.00	63770.43

Si considera un dataset “insurance”, un campione che detiene campi come: età, sesso, indice corporeo, numero di figli, fumatori (sì, no), regione di appartenenza e spesa sanitaria (age, sex, bmi, children, smoker, region, charges). Il campione è rappresentato a parità di genere.

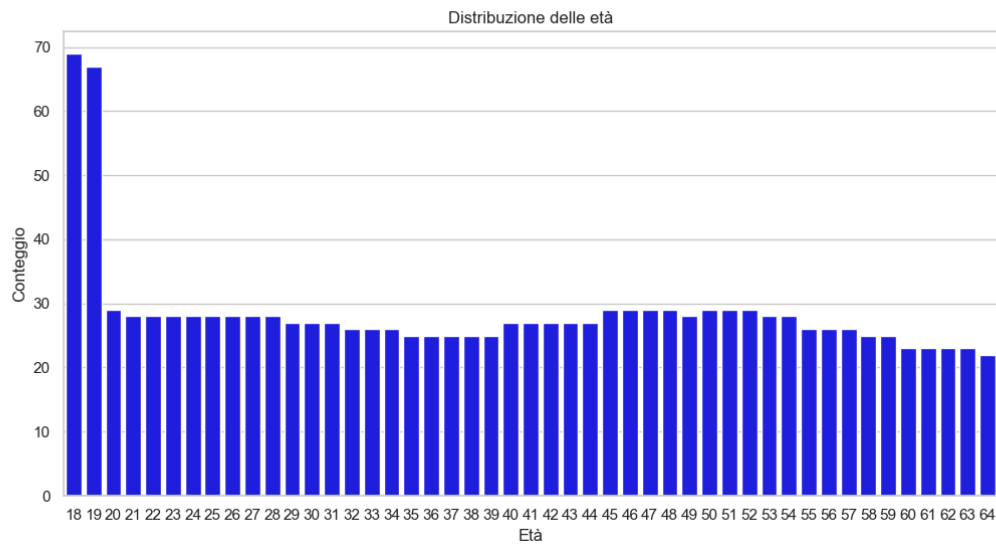
In primo luogo si prendono in considerazione le analisi statistiche (media, moda, mediana, minimo, massimo, primo e terzo quartile)

(Fonte: <https://www.kaggle.com/datasets/mirichoi0218/insurance>)

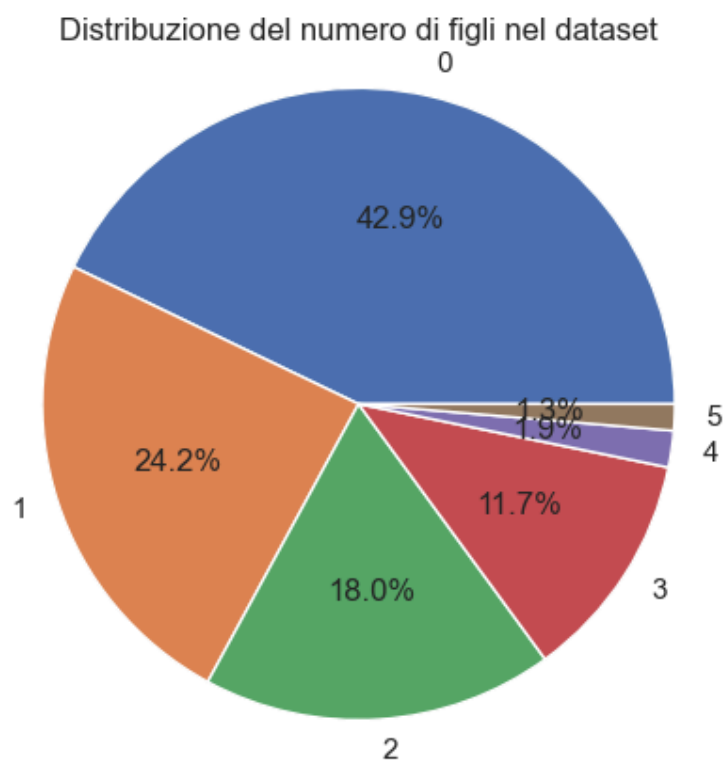
Il dataset prende in considerazione una popolazione da un'età compresa tra i 18 e 64 anni, la divisione in percentuale tra la popolazione maschile e femminile è rispettivamente 50,5% e 49,5%.

La percentuale dei fumatori risulta inferiore (20,5%) rispetto ai non fumatori (79,5%).

Analisi di coorte

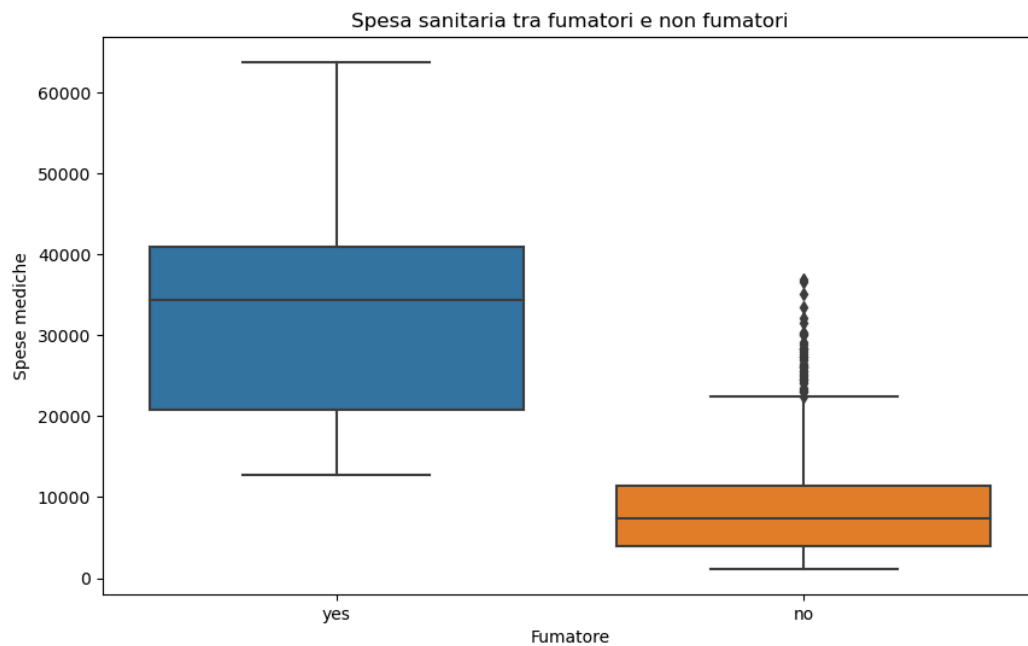


Facendo un'analisi approfondita di come è distribuito il dataset, è emerso che in realtà è rappresentato da persone che hanno diciotto anni, anche se la media è di 39 anni, quindi influenzata da degli outlier.



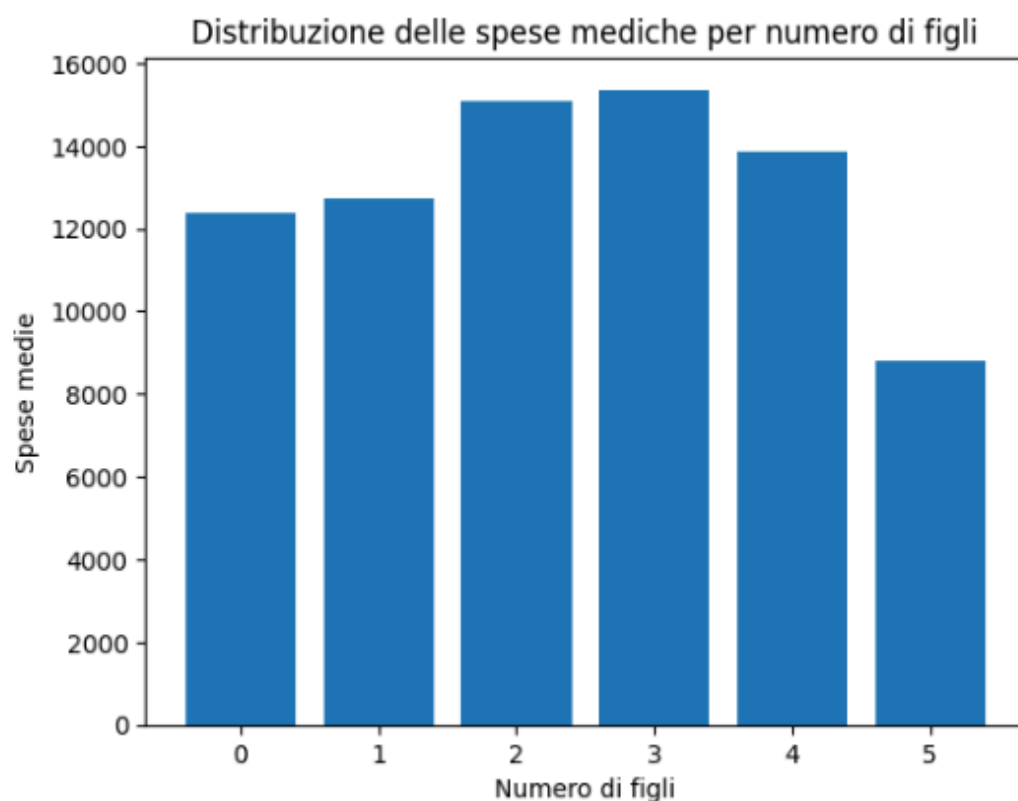
La maggior parte della popolazione (42,9%) non ha figli a carico, mentre il 24% della popolazione ha solo un figlio a carico.

Spesa media tra fumatori e non fumatori

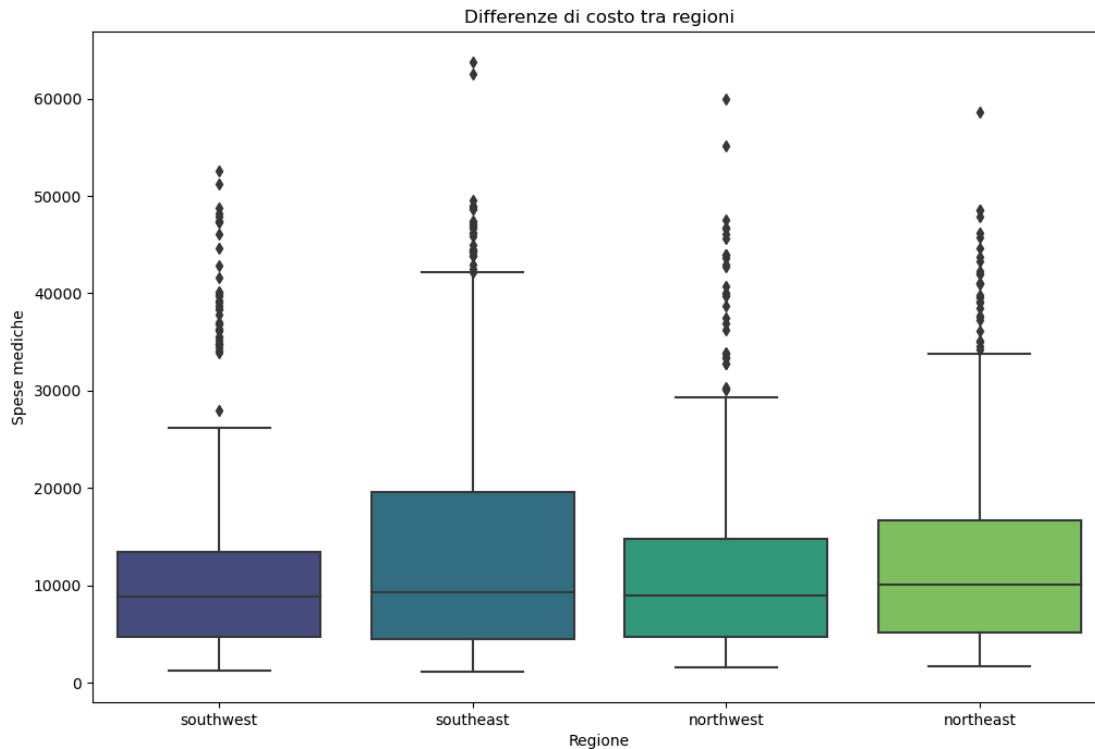


Le spese medie dei fumatori (circa 8440 \$) sono significativamente più alte rispetto a quelle dei non fumatori (circa 32050 \$), con una differenza evidente. La distribuzione delle spese mediche dei non fumatori è distribuita in un range di spesa inferiore rispetto ai fumatori.

Significativo è il fatto che considerando il dataset che è rappresentato da quasi l'80% per cento da non fumatori, rappresenta che le due variabili sono direttamente proporzionali.

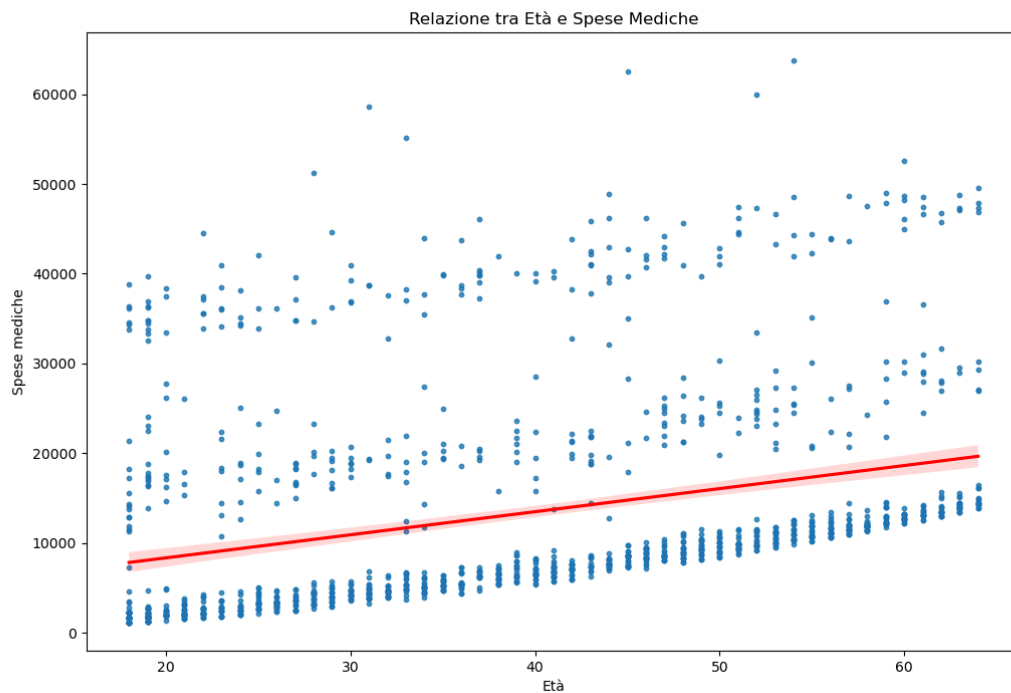


Questo grafico rappresenta la spesa media sanitaria in base al numero di figli a carico. Per le persone che hanno tra l'uno o due figli si ha maggiore spesa sanitaria. Chi non ha figli ha quasi le stesse spese mediche per chi detiene un solo figlio. Da più di un figlio, la spesa media si alza ed è quella che richiede maggior spesa stessa. A un certo punto però la spesa scende dal terzo figlio in poi, ciò suggerisce che un maggior numero di figli minore sarà la spesa sanitaria.



Altre variabili prese in esame sono la distribuzione della spesa sanitaria tra le diverse regioni considerate (South-west, South-east, Nord-west, North-east), per capire se c'è o meno differenziazione.

Non ci sono particolari differenze nella spesa sanitaria nelle diverse regioni. Nel dettaglio si evince che il range di spesa è maggiore nella regione Southeast rispetto alle altre regioni.



Questo grafico rappresenta la correlazione tra spesa medica ed età. Si può notare come sembra che ci siano tre fasce di spesa ben definite suggerendo tre piani di assicurazione sanitaria. La prima fascia, quella più economica, sembrerebbe anche quella più rappresentata. Notiamo, inoltre, una correlazione positiva tra l'età e la spesa medica.

Analisi per condizione fisica (BMI)

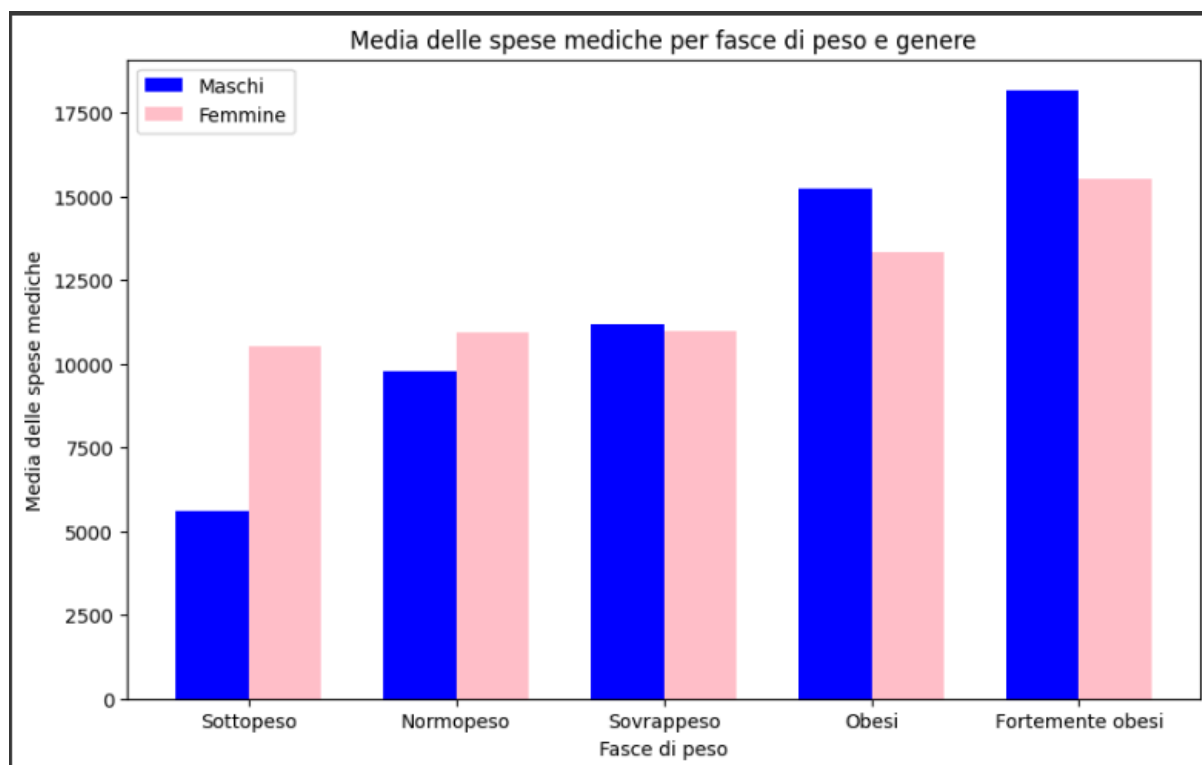
In questa sezione, sono state esplorate le possibili correlazioni tra condizione fisica legata all'indice di massa corporea (BMI) e le spese mediche. L'obiettivo è comprendere se esiste una relazione tra l'indice di massa corporea (BMI) e le spese mediche.

Ponendo come :

- Sottopeso: BMI < 18,5
- Normopeso: 18,5 < BMI < 24,9
- Sovrappeso: 25 < BMI < 29,9
- Fortemente Sovrappeso: 30 < BMI < 34,9
- Severamente Obese: 35 < BMI



In questi grafici a torta sono state suddivise le BMI per persone Sottopeso, Normopeso, Sovrappeso, Obese e Fortemente, dividendo ancora per genere. Ciò che emerge è che nel caso del sottopeso e normopeso questa categoria è rappresentata più da donne. Nel caso delle restanti categorie gli uomini sono un po' di più delle donne.



In questo caso questo grafico descrive la spesa media ripartita secondo la categoria di BMI. Si nota generalmente che la spesa le categorie che pesano di più sulla spesa sanitaria sono coloro che hanno una forma di obesità grave. In maniera opposta si considera come nel caso del sottopeso la spesa sanitaria sia più alta rispetto a un normopeso e che è influenzata in particolar modo dal sesso femminile. Anche nel caso del normopeso le donne hanno una media leggermente più alta sulla spesa sanitaria. Poi c'è un'inversione di tendenza, dove emerge che gli uomini superano le donne nella spesa sanitaria man mano che l'ibm diventa diverso cioè più l'obesità si aggrava più la spesa aumenta (soprattutto nel caso degli uomini)

Gruppo Horror

Collaboratori:
 Albano Martina
 Branca Federica
 Cerri Marco Renato
 Mininni Claudia
 Stefanetti Margot