

#### **Caso Studio:**

Durante la pandemia COVID-19 molti lavoratori hanno subito in generale una variazione importante della propria routine sonno-lavoro, quindi sono ipotizzabili degli effetti nocivi nella quotidianità di molti lavoratori anche dopo la conclusione della pandemia.

Si vuole realizzare un'analisi interna per gli Operatori di Esercizio di GTT per comprendere quali siano i fattori associati all'aumento del numero di incidenti e il peggioramento della qualità del sonno. Si è svolto un sondaggio online anonimo al sopracitato personale GTT, dal quale abbiamo ottenuto informazioni sulle ore di sonno al giorno, classifica della qualità del sonno espressa tra 1 e 10, eventuale consumo di alcol, tempo dedicato alla attività fisica, numero di sigarette al giorno se fumatori ed infine tempo trascorso davanti ad uno schermo. Inoltre come ulteriore fonte di dati per il caso studio si sono utilizzati Open Data presi da ISTAT e da GTT stessa.

### **Dataset**

Il primo dataset a disposizione è stato prodotto in formato "<u>JSON</u>", questi dati sono stati raccolti tramite un sondaggio online eseguito dai dipendenti GTT. Questo dataset è stato in seguito filtrato a dipendenti aventi una specifica qualifica, in questo caso gli Operatori di Esercizio. In particolare sono stati ottenuti per ogni soggetto le ore di sonno e la sua classifica espressa da 1 (pessimo) a 10 (ottimo), l'eventuale uso di alcol con commento annesso sulla quantità approssimativa, la quantità di attività fisica eseguita regolarmente in settimana, le sigarette fumate quotidianamente se fumatori, la quantità di ore trascorse davanti allo schermo di un qualsiasi dispositivo elettronico ed infine un feedback personale su come si sentono.

Il secondo dataset scelto è stato prodotto in formato "<u>XML</u>", questi dati sono stati raccolti da ACI e forniti al pubblico sotto licenza <u>IODL v2.0</u>. Tale dataset è stato opportunamente rielaborato in modo da avere solo gli incidenti in cui un mezzo della GTT sia stato coinvolto. In particolare sono stati ottenuti il luogo dell'incidente (Comune, Provincia, Regione) e la gravità dell'incidente quantificato secondo numero di morti o feriti.

Il terzo dataset scelto è stato prodotto in formato ".CSV", questi dati sono stati raccolti da GTT e concessi per l'analisi interna dell'azienda e quindi non pubblici. Tale dataset avrà la seguente struttura: numero identità del dipendente, anno di riferimento, genere del dipendente, qualifica del dipendente, orario dei turni di lavoro, linea del trasporto pubblico a cui sono assegnati ed il tipo di mezzo guidato.

# Analisi e raccolta dei requisiti

Si vuole eseguire per conto del Gruppo Torinese Trasporti un'analisi sugli effetti nella qualità del sonno in seguito al *COVID-19* e come questo influisca sull'aumento di *incidenti stradali* del proprio parco autovetture. Si desidera inoltre individuare se altre abitudini quali alcol, fumo e l'eccessivo tempo davanti allo schermo possano influire in modo significativo ad una diminuzione dell'indice di attenzione durante l'orario di lavoro, principale causa degli incidenti su strada.

Al fine di eseguire tali analisi saranno proposte le seguenti statistiche: la media delle *ore di sonno* (espresse in ore) per *turno di lavoro*, la media della qualità del sonno per turno di lavoro, la media delle sigarette consumate per attività fisica, ed infine il numero medio di ore che i soggetti trascorrono davanti uno schermo per qualità del sonno.

## Strutturazione dei requisiti

**Frase di carattere generale:** L'azienda GTT (Gruppo Torinese Trasporti) ha richiesto un'analisi ad uso interno nel quale si valutino la variazione della qualità del sonno nel periodo successivo al COVID-19 e se esista una relazione tra l'aumento di incidenti stradali dei propri mezzi con eventuali cattive consuetudini acquisite durante il periodo pandemico.

**Frase di carattere specifico:** Il cliente GTT richiede, al fine di valutare il rendimento e lo stato di salute dei propri Operatori di Esercizio, uno studio sulle abitudini acquisite durante il periodo pandemico; in particolare si presterà attenzione all'uso di alcol, fumo ed all'uso eccessivo di dispositivi elettronici con schermi luminosi. Inoltre si studierà quante ore i dipendenti riescano e dormire a seconda della fascia oraria del proprio turno di lavoro.

### Glossario

Termine	Descrizione	Sinonimi
Incidenti	Uno scontro che avviene su una strada pubblica, che coinvolge	Sinistro
stradali	almeno un veicolo e che può conseguenze sulla salute di chi vi è	stradale
	coinvolto.	
Abitudini	Tendenza alla ripetizione di un determinato comportamento,	Consuetudine
	collegabile a fattori naturali o acquisiti.	
Ore di sonno	Necessità fisiologica che si contrappone allo stato di veglia	Dormita
Turno di	Metodo di organizzazione del lavoro secondo il quale l'orario del	Rotazione
lavoro	lavoro segue un determinato ritmo.	
COVID-19	Sindrome respiratoria acuta grave.	Malattia

#### Ruoli

- Il Data Engineer ha avuto il ruolo di trovare i database.
- Il Data Analyst ha avuto il ruolo di unire i dataset e completare l'analisi di business.
- Il Data Scientist ha avuto il ruolo di compiere le analisi predittive.