

Progetti

Angelica Lo Duca
angelica.loduca@iit.cnr.it

Esame - Progetto di Data Journalism

- Lavoro di gruppo da 1 a 3 persone max
- Scelta del progetto concordata
- Sottomettere il codice e i dati
- L'esame orale è una discussione del progetto

Quest'anno suggeriamo noi dei temi ma potete proporre anche voi delle idee.

5 macro temi

1. Società dell'informazione

Uso di Internet in Italia

```
graph TD; A["Uso di Internet in Italia"] --> B["Audiweb  
Audience online degli editori iscritti al servizio."]; A --> C["Fonte Istat  
Accesso e tipo di utilizzo"];
```

Audiweb

Audience online degli editori iscritti al servizio.

Fonte Istat

Accesso e tipo di utilizzo

Audiweb

*Dati organizzati per anno:
tanti dataset, uno per
ogni anno.*

Data Cleaning

```
graph LR; A["Audiweb  
Dati organizzati per anno:  
tanti dataset, uno per  
ogni anno."] --> B["Data Cleaning"]
```

Audiweb: Possibili domande per estrarre un insight

1. Come è cambiato nel tempo l'interesse dell'audience?
2. Qual è stato l'interesse maggiore in un periodo specifico (e.g. Covid-19, mondiali di calcio, guerra in Ucraina, ...)?
3. Come varia l'interesse in base alle diverse fasce di età?
4. Ci sono differenze significative nel consumo digitale tra utenti giovani (18-34 anni) e utenti più anziani?
5. Quali sono le testate o i servizi digitali che hanno avuto la crescita più significativa rispetto al mese precedente?

Istat: Possibili domande

1. Qual è stata l'evoluzione della percentuale di famiglie con accesso a Internet negli anni disponibili?
2. Quale tipo di connessione a Internet è più diffuso tra le famiglie?
3. Come varia la disponibilità di Internet tra diverse tipologie familiari?
4. C'è una tendenza di crescita o calo nell'uso della banda larga fissa rispetto alla banda larga mobile?
5. Quali sono i periodi di maggiore variazione nell'accesso a Internet da parte delle famiglie?

Procedere nell'esplorazione dei dati
finché non si trova qualcosa di
interessante (l'insight).

Una volta trovato l'insight,
approfondire l'argomento con dati
provenienti da altre sorgenti.

Possibili estensioni

Ambito economico

- Diffusione dello smart working
- Percentuale di popolazione che effettua acquisti online

Ambito sanitario

- Uso di telemedicina
- Impatto di Internet sulla salute mentale

Ambito educativo

- Correlazione tra connessione ad Internet e rendimento scolastico

Esempio di insight: durante il periodo del Covid-19, il brand X ha ottenuto un incremento di visite di Y% e il numero di famiglie con accesso ad Internet è aumentato di Z%.

Durante il Covid-19 sono anche aumentati i rischi relativi alla salute mentale?

Il rendimento scolastico migliora con l'accesso ad Internet?

Con l'introduzione dello smart working è aumentato il numero di visite per un brand?

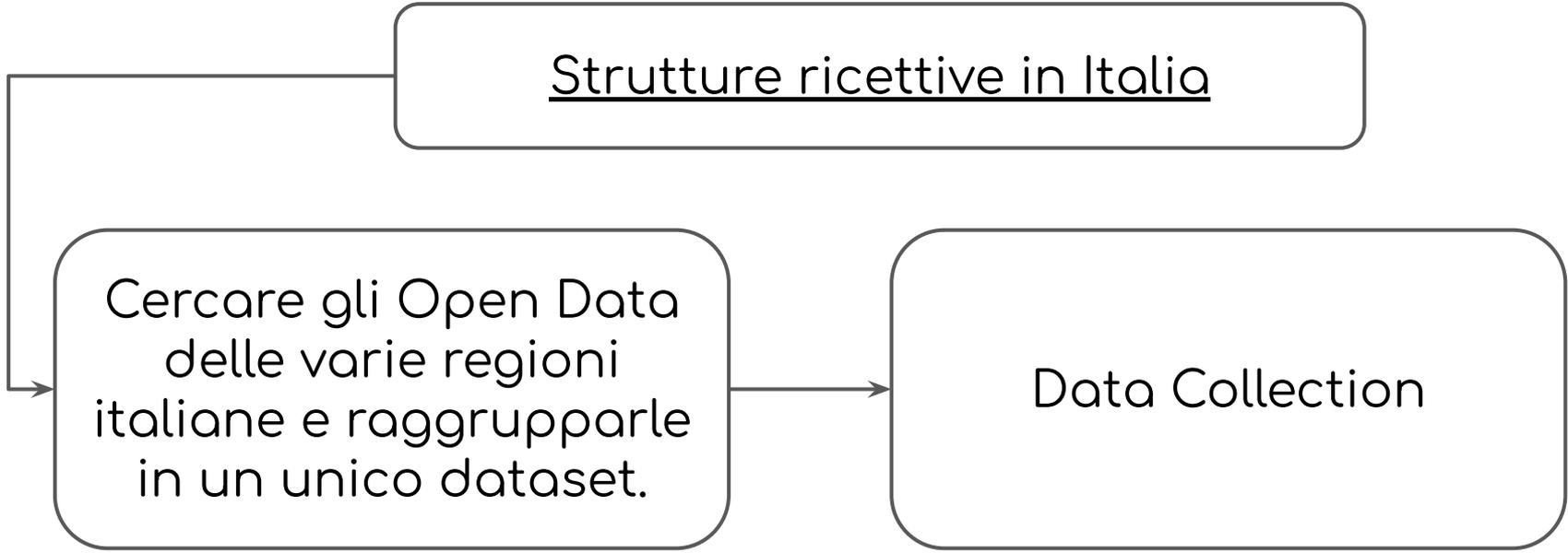
L'insight descritto è di tipo temporale perché analizza l'evoluzione nel tempo.

2. Turismo

Strutture ricettive in Italia

Cercare gli Open Data
delle varie regioni
italiane e raggrupparle
in un unico dataset.

Data Collection



Possibili domande

1. Quali sono le regioni con il maggior numero di strutture ricettive?
2. Quali sono le regioni con la maggiore capacità ricettiva (numero di posti letto)?
3. Qual è il rapporto tra numero di strutture e popolazione residente per ogni regione?
4. Quali regioni hanno registrato il maggior aumento/diminuzione delle strutture negli ultimi anni?
5. Esiste un divario tra Nord e Sud, anche in termini di numero di stelle delle strutture ricettive?

Procedere nell'esplorazione dei dati
finché non si trova qualcosa di
interessante (l'insight).

Una volta trovato l'insight,
approfondire l'argomento con dati
provenienti da altre sorgenti.

Possibili estensioni

Ambito economico

- Come varia il costo medio del pernottamento in base alla stagione e alla regione

Ambito demografico

- Rapporto tra turismo e popolazione residente

Ambito medico

- Incidenza delle malattie stagionali legate al turismo

Esempio di insight: la regione X ha il più alto tasso di strutture per km² ma le strutture contenute hanno pochi posti letto.

Come variano i costi nella regione X rispetto alle altre regioni?

La popolazione residente nella regione X accetta benevolmente il turismo?

Come cambia l'accesso ai pronto soccorso nella Regione X durante l'alta stagione?

L'insight descritto è di tipo spaziale perché analizza come è distribuita una risorsa nello spazio.

3. Telecomunicazioni globali

Cavi sottomarini

Cercare i dati e
raggrupparli in un
unico dataset.

Data Collection
Ad esempio:

<https://www.submarine-cablemap.com/>

Possibili domande

1. Quali possibili percorsi può seguire un segnale che va dall'Italia alla Cina?
2. Come comunicano Cina e Stati Uniti?
3. Negli anni di quanto è variato il numero di cavi sottomarini?
4. Quanti sono i cavi non utilizzati?
5. Quali sono le aree meno servite?

Procedere nell'esplorazione dei dati
finché non si trova qualcosa di
interessante (l'insight).

Una volta trovato l'insight,
approfondire l'argomento con dati
provenienti da altre sorgenti.

Possibili estensioni

Ambito geo-politico

- Le infrastrutture dei cavi sottomarini sono più vulnerabili in alcune aree geopoliticamente instabili?

Ambito economico

- I paesi con più connessioni sottomarine hanno un vantaggio nel commercio digitale e nei servizi finanziari?

Ambito ambientale

- I cavi sottomarini hanno un impatto sugli ecosistemi marini e sulla biodiversità?

Esempio di insight: In corrispondenza del cavo X si verifica un'anomalia di tipo Y.

Rispetto al cavo X, come sono posizionate le potenze politiche?

Cosa accade a livello ambientale nei pressi del cavo X?

Qual è l'impatto economico dell'anomalia generata dal cavo X?

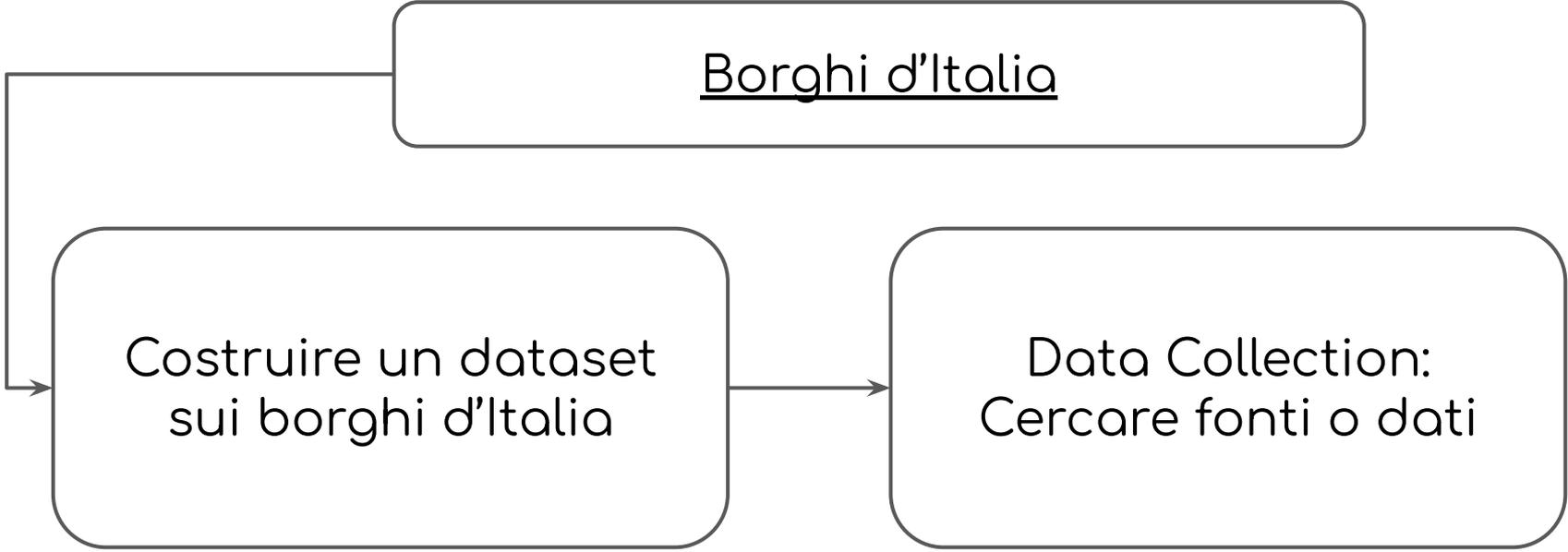
L'insight descritto è di tipo anomalo perché analizza un'anomalia presente nei dati.

4. Patrimonio Culturale

Borghi d'Italia

Costruire un dataset
sui borghi d'Italia

Data Collection:
Cercare fonti o dati



Possibili domande

1. In quali regioni d'Italia sono maggiormente concentrati borghi?
2. Qual è il numero minimo, medio e massimo di abitanti per borgo?
3. Come si è evoluta la popolazione dei borghi nel tempo?
4. Qual è la distanza dei borghi dai centri?

Procedere nell'esplorazione dei dati
finché non si trova qualcosa di
interessante (l'insight).

Una volta trovato l'insight,
approfondire l'argomento con dati
provenienti da altre sorgenti.

Possibili estensioni

Ambito demografico

- Quali borghi stanno subendo lo spopolamento più velocemente?

Ambito economico

- Qual è l'impatto degli incentivi economici sullo sviluppo dei borghi?

Ambito Infrastrutturale e Mobilità

- I borghi con un migliore accesso ai trasporti hanno meno problemi di spopolamento?

Esempio di insight: I borghi con una migliore accessibilità ai trasporti pubblici e alla rete stradale tendono a mantenere più stabile la popolazione rispetto a quelli più isolati.

I borghi con una maggiore accessibilità ai trasporti hanno maggiore accesso ad incentivi?

Come è variata la popolazione dei borghi più isolati nel tempo?

Qual è l'impatto del turismo nei borghi più isolati?

L'insight descritto è di tipo relazionale perché fa un confronto tra diverse categorie di dati.

5. Salute

Incidenza del Covid-19

```
graph TD; A[Incidenza del Covid-19] --> B[Morti durante il periodo del Covid in Italia (Istat)]; A --> C[Morti durante il periodo del Covid in Francia/Europa];
```

Morti durante il
periodo del Covid
in Italia (Istat)

Morti durante il
periodo del Covid
in Francia/Europa

Possibili domande

1. Di quanto sono aumentati i morti durante il periodo del Covid?
2. Si registra una differenza sostanziale tra l'Italia e il resto dell'Europa?
3. Quanti ricoverati in terapia intensiva sono guariti?

Procedere nell'esplorazione dei dati
finché non si trova qualcosa di
interessante (l'insight).

Una volta trovato l'insight,
approfondire l'argomento con dati
provenienti da altre sorgenti.

Possibili estensioni

Ambito demografico

- Di quanto è diminuita la popolazione nelle città che hanno subito più morti per Covid-19?

Ambito economico

- Qual è l'impatto economico delle morti per Covid-19 nei piccoli centri?

Ambito dei trasporti

- Come è variata la viabilità durante il periodo del Covid-19?

Esempio di insight: Il comune X ha avuto il minor numero di morti per Covid-19 in percentuale, mentre il comune Y ha avuto il maggior numero di morti.

Quali misure relative alla viabilità ha adottato il comune X?

Quali strategie economiche ha adottato il comune X durante il Covid-19? E il comune Y? E gli altri comuni?

Di quanto è diminuita la popolazione nei comuni X e Y?

L'insight descritto è di tipo scalare perché analizza un fenomeno dal piccolo al grande.

Tematiche progetti

- Società dell'informazione
 - Uso di Internet in Italia
- Turismo
 - Strutture ricettive in Italia
- Telecomunicazioni globali
 - Cavi sottomarini per telegrafo, Internet
- Patrimonio culturale
 - Borghi d'Italia
- Salute
 - Morti in Italia, Francia, Europa durante il periodo del Covid-19