

REPORT PER IL PROGETTO: «Indice di Qualità (IQA) nell'agglomerato di Milano: analisi dell'andamento stagionale e annuale»

Progetto di:

Gianluca CAVALLARO matricola n. 826049

Gianluca SCURI matricola n. 886725

Giorgio CARBONE matricola n. 811974

Remo MARCONZINI matricola n. 883256



INTRODUZIONE AL PROBLEMA, AL DOMINIO E ALLA DOMANDA DI RICERCA

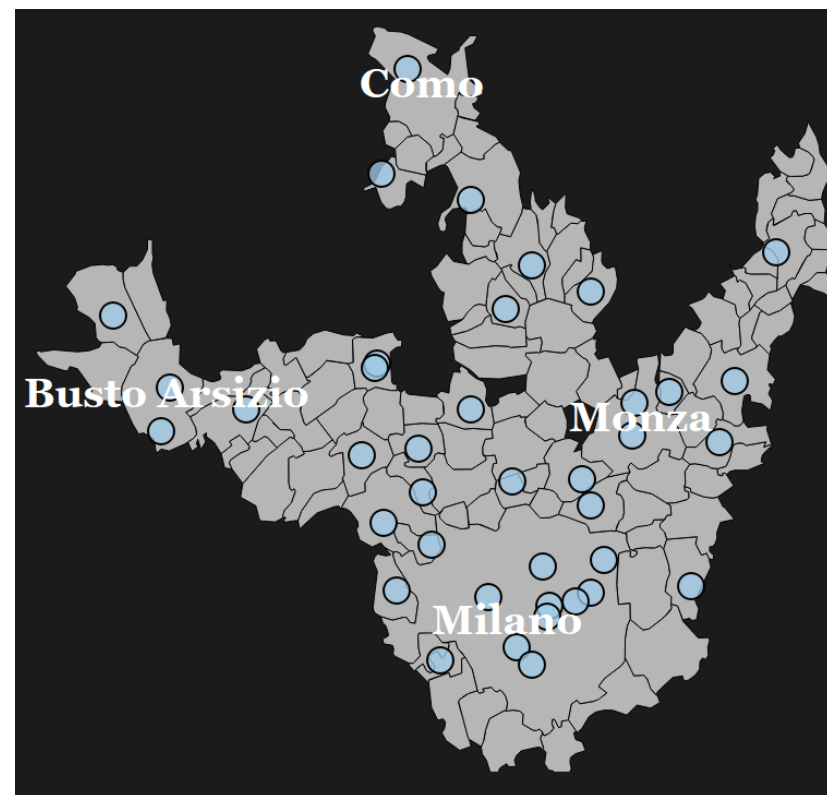
Cos'è l'Indice di Qualità dell'Aria (IQA) ?

- ❑ L'esposizione prolungata ad aria inquinata può causare **problemi respiratori e cardiovascolari**
- ❑ Per questo è **importante monitorare** i livelli di inquinamento atmosferico
- ❑ L'**Indice di Qualità dell'Aria (IQA)** è un indicatore sintetico per la definizione della qualità dell'aria ed è calcolato sulla base dei livelli dei 5 principali inquinanti: **PM10, PM2.5, O3, NO2 e SO2**
- ❑ La qualità dell'aria è valutata in funzione di ciascuno dei 5 inquinanti, sulla base dei limiti europei
- ❑ L'**IQA complessivo** dipende, però, dall'**inquinante** che ottiene la **valutazione peggiore** e definisce la qualità dell'aria complessiva come:
Molto Buona, Buona, Accettabile, Scarsa, Molto Scarsa

Monitoraggio dell'aria nell'agglomerato di Milano

- ❑ **ARPA Lombardia** effettua la **valutazione della qualità dell'aria** in maniera sistematica e costante, sull'intero territorio regionale (secondo la Direttiva 2008/50/CE, recepita dal D. Lgs. 155/2010), attraverso una rete di stazioni diffuse sul territorio
- ❑ I dati ottenuti dalle rilevazioni effettuate da ARPA sono pubblicati sulla piattaforma [Open Data di Regione Lombardia](#)
- ❑ Il territorio lombardo è suddiviso in zone e agglomerati urbani
- ❑ I territori della Val Padana, per via dell'alto tasso di urbanizzazione e di fattori orografici e meteorologici, sono tra le zone più inquinate d'Europa
- ❑ L'agglomerato di Milano include circa 3 milioni di abitanti, comprende le città di Milano, Monza, Como e territori limitrofi e viene costantemente monitorato attraverso **39 stazioni**

Mappa dei comuni dell'agglomerato di Milano e posizione delle stazioni di monitoraggio



Domande di ricerca

- ❑ Qual è stato l'impatto dell'implementazione delle misure legislative e tecnologiche Europee sulla qualità dell'aria nell'agglomerato di Milano dal 2006 a oggi ?
- ❑ Come è caratterizzato l'andamento stagionale della qualità dell'aria nell'agglomerato di milano ?
 - ❑ Quali sono i fattori meteorologici e antropici che ne determinano l'andamento ?
- ❑ Qual'è stato l'andamento dei livelli dei singoli inquinanti principali, e più pericolosi per la salute umana e l'ambiente, dal 2006 a oggi ?
 - ❑ Quali sono le fonti di emissioni antropogeniche maggiori ?
 - ❑ Quali sono i fattori meteorologici e antropici che influenzano l'andamento stagionale dei singoli inquinanti ?

I DATI: FONTI, *TOOLS*, PREPARAZIONE E ANALISI

Tools utilizzati per la preparazione e l'analisi

❑ Linguaggio di programmazione e piattaforma utilizzata:

- [Python](#): linguaggio di programmazione
- [Jupyter notebook](#): applicazione web open-source per il calcolo interattivo

❑ Librerie utilizzate:

- [Pandas](#): libreria open-source per l'analisi e la manipolazione dei dati
- [Numpy](#): libreria per il calcolo scientifico
- [Seaborn](#): libreria per la visualizzazione dei dati

❑ Altri software:

- [Tabula](#): software gratuito per lo scraping da file PDF

Dataset relativi alla qualità dell'aria: descrizione e fonti

❑ [Dati sensori aria](#) riporta i **valori misurati dai sensori di rilevazione della concentrazione degli inquinanti** dal 1996 al 2021, in particolare contiene:

- Rilevazione di ogni sensore per ogni istante (IdSensore, Valore, Data) con frequenza oraria (NO2, SO2, O3) e giornaliera (PM)
- Indicazioni sulla validità del dato (dato) e se si tratta di un dato aggregato (idOperatore)

❑ [Stazioni qualità dell'aria](#) contiene uno storico dei **sensori installati negli anni** sul suolo lombardo

- Attributi che identificano il sensore (IdSensore, NomeTipoSensore, UnitàMisura)
- Attributi che identificano la stazione (IdStazione, NomeStazione, DataStart, DataStop)
- Attributi di localizzazione (Provincia, Comune)

	IdSensore	Data	Valore	Stato	idOperatore
0	5504	1996-01-01 00:00:00	61.0	VA	1
1	5505	1996-01-01 00:00:00	61.0	VA	1
2	5506	1996-01-01 00:00:00	53.9	VA	1
...
98291919	10331	2021-12-20 18:00:00	27.6	VA	1
98291920	10333	2021-12-25 11:00:00	84.7	VA	1
98291921	10331	2021-12-21 04:00:00	34.3	VA	1

98291922 rows × 5 columns

	IdSensore	NomeTipoSensore	UnitàMisura	Idstazione	NomeStazione	Quota	Provincia	Comune	Storico	Utm_Nord
0	17286	Biossido di Azoto	µg/m³	1374	Monza Parco	181	MB	Monza	N	5049780
1	17284	Ammoniaci	µg/m³	1374	Monza Parco	181	MB	Monza	N	5049780
2	17290	PM10 (SM2005)	µg/m³	1374	Monza Parco	181	MB	Monza	N	5049780
3	17285	Ossidi di Azoto	µg/m³	1374	Monza Parco	181	MB	Monza	N	5049780
4	17288	Ozono	µg/m³	1374	Monza Parco	181	MB	Monza	N	5049780
...
961	6606	Particolato Totale Sospeso	µg/m³	591	Seriate v. Garibaldi	256	BG	Seriate	S	5059449
962	6382	Ossidi di Azoto	µg/m³	569	Sondrio v. Mazzini	307	SO	Sondrio	N	5113078
963	6607	Particolato Totale Sospeso	µg/m³	591	Seriate v. Garibaldi	256	BG	Seriate	S	5059449
964	20523	Ammoniaci	µg/m³	583	Bergamo v. Meucci	249	BG	Bergamo	N	5059922
965	12597	Cadmio	ng/m³	609	Casirate d'Adda v. Cimitero	108	BG	Casirate d'Adda	N	5038450

966 rows × 10 columns

Dataset relativi alla qualità dell'aria: preparazione dei dati

❑ Dataset rilevazioni sensori qualità dell'aria:

- Rimozione righe con valori mancanti (~11% delle righe totali)
- Rimozione righe con IdOperatore non valido
- Risultato: dataset composto da 87 milioni di righe

❑ Dataset lista dei sensori:

- Rimozione sensori di grandezze diverse da PM2.5, PM10, NO2, SO2 e O3 (gli inquinati considerati nel calcolo dell'IQA)

Dataset relativi alla qualità dell'aria: analisi dei dati

- ❑ **Raggruppamento delle righe del dataset** delle misurazioni rispetto al **singolo sensore** e al **giorno** e **calcolo delle grandezze aggregate**
 - Calcolo della **media**, del **massimo**, dei **conteggi** (numero di valori mediati) e della **deviazione standard**
- ❑ **Unione** dei due dataset (rilevazioni e lista dei sensori) rispetto **all'Id del sensore** (attributo usato come key del JOIN)
 - Conseguente perdita delle rilevazioni di 88 sensori poiché non sono presenti nel dataset lista dei sensori
- ❑ **Verifica** validità delle medie ottenute
 - Le rilevazioni dei sensori del **PM2.5** e del **PM10** sono **medie giornaliere** ed hanno un conteggio pari a 1
 - Per gli altri sensori vengono **rimosse le medie giornaliere** ottenute da **meno di 8 osservazioni** che corrispondono a 8 ore
- ❑ **Scelta della valore aggregato corretto** per ogni tipo di sensore rispetto al valore richiesto nel calcolo dell'indice di qualità dell'aria
 - Per **PM2.5** e **PM10** tengo il valore medio giornaliero -> La qualità dell'aria in funzione di PM è valutata in termini di concentrazione media giornaliera
 - Per **NO2**, **SO2** e **O3** tengo il valore massimo giornaliero -> La qualità dell'aria in funzione di questi inquinanti è valutata in termini di concentrazione oraria massima giornaliera

Dataset relativi alla zonizzazione: descrizione e fonti

❑ [Tabella zonizzazione del territorio lombardo](#) è un file PDF contenente la **lista di tutti i comuni lombardi** con associata un **zona di appartenenza**

- L'attributo **Comune** è il nome della municipalità
- L'attributo **Zona** può assumere i valori: A, B, C, D, Agglomerato di Milano, Agglomerato di Bergamo, Agglomerato di Brescia

1. Classificazione dei Comuni del territorio lombardo all'interno degli agglomerati e delle zone A, B, C e D in relazione a tutti gli inquinanti, ad esclusione dell'ozono.

Provincia	Codice Istat	Comune	Zona	Residenti 2008	Superficie (ha)
BG	16009	AMBIVERE	A	2341	327
BG	16013	ARZAGO D'ADDA	A	2836	944
BG	16018	BAGNATICA	A	4119	639
BG	16020	BARIANO	A	4396	714
BG	16021	BARZANO'	A	5178	356
BG	16028	BOLGARE	A	5538	859
BG	16030	BONATE SOPRA	A	8637	605
BG	16031	BONATE SOTTO	A	6538	634
BG	16034	BOTTANUCO	A	5243	581
BG	16038	BREMBATE DI SOPRA	A	7768	439
BG	16040	BRIGNANO GERA D'ADDA	A	5798	1201
BG	16043	CALCINATE	A	5791	1517

Dataset relativi alla zonizzazione: preparazione dei dati

- ❑ Creazione di un **file csv a partire dal file PDF** utilizzando il software Tabula

	Provincia	Codice Istat	Comune	Zona	Residenti 2008	Superficie (ha)
0	BG	16009	AMBIVERE	A	2341	327
1	BG	16013	ARZAGO D'ADDA	A	2836	944
2	BG	16018	BAGNATICA	A	4119	639
...
1542	VA	12132	VARANO BORGHI	C	2313	328
1543	VA	12135	VEDDASCA	C	290	1701
1544	VA	12139	VIGGIU'	C	5243	928

1545 rows × 6 columns

Dataset relativi alla zonizzazione: analisi dei dati

- ❑ **Unione** del dataset **zonizzazione** con il **dataset ottenuto precedentemente** dall'unione dei valori dei sensori della qualità dell'aria con le relative stazioni
 - Unione rispetto all'**attributo testuale Comune**
- ❑ **Raggruppamento delle righe** del dataset appena ottenuto rispetto al singolo sensore, al giorno e alla zona, e calcolo delle **grandezze aggregate**
 - Calcolo della **media**, dei **conteggi** e della **deviazione standard**
- ❑ **Subsetting**
 - **Rimozione** delle righe corrispondenti ad osservazioni di **stazioni ubicate fuori** dalla zona **Agglomerato di Milano**
 - Decidiamo di **non considerare** gli anni **dal 1996 al 2005** poiché non erano presenti sensori per la rilevazione individuale di PM2.5 e PM10
- ❑ **Calcolo dell'indice di qualità** dell'aria relativa al **singolo inquinante** considerando gli [intervalli tabulati](#) da ARPA Lombardia e calcolo dell'indice di qualità **complessivo** prendendo il valore dell'inquinante peggiore in quel giorno

Dataset relativi al meteo: descrizione e fonti

❑ [Dati sensori meteorologici](#) riporta i **valori misurati dai sensori** dal 2017 al 2021

- Rilevazione di ogni sensore con frequenza di 10 minuti
- Indicazioni sulla validità del dato e se si tratta di un dato aggregato (Stato, idOperatore)

❑ [Stazioni meteorologiche](#) contiene uno storico dei **sensori installati negli anni** sul suolo lombardo

- Attributi che identificano il sensore (IdSensore, Tipologia, Unità DiMisura)
- Attributi che identificano la stazione (IdStazione, NomeStazione)
- Attributi di localizzazione (Provincia, Ing, lat)

	IdSensore	Valore	Stato	idOperatore	Data_corretta
2	10377	1.4	VA	1	2017-01-01
3	10381	81.6	VA	1	2017-01-01
4	10382	3.2	VA	1	2017-01-01
7	10551	-2.6	VA	1	2017-01-01
8	10552	100.0	VA	1	2017-01-01

288126365 rows × 5 columns

	IdSensore	Tipologia	Unità DiMisura	IdStazione	NomeStazione	Quota	Provincia	DataStart	DataStop	Storico	UTM_Nord	UTM_Est
0	22006	Umidità Relativa	%	1890	Tavernole sul Mella Monte Guglielmo	1790	BS	14/10/2019	NaN	N	5067376	591417
1	22007	Direzione Vento	°	1890	Tavernole sul Mella Monte Guglielmo	1790	BS	14/10/2019	NaN	N	5067376	591417
2	22003	Temperatura	°C	1890	Tavernole sul Mella Monte Guglielmo	1790	BS	14/10/2019	NaN	N	5067376	591417
...
1259	19394	Direzione Vento	°	1650	Bormio eliporto	1172	SO	29/09/2016	20/10/2020	S	5145365	604906
1260	12757	Temperatura	°C	1511	Rescaldina	215	MI	25/11/2020	NaN	N	5050479	496327
1261	19393	Velocità Vento	m/s	1650	Bormio eliporto	1172	SO	29/09/2016	20/10/2020	S	5145365	604906

1262 rows × 15 columns

Dataset relativi al meteo: preparazione dei dati

❑ Dataset rilevazioni:

- Rimozione righe con stato nullo
- Rimozione righe con IdOperatore non valido
- Risultato: dataset composto da 246 milioni di righe

❑ Dataset sensori:

- Rimozione sensori di grandezze diverse da temperatura e radiazione globale

Dataset relativi al meteo: analisi dei dati

- ❑ Raggruppamento del dataset dei valori, rispetto al singolo sensore e al giorno, e calcolo delle **grandezze aggregate**
 - Calcolo della **media**, del **massimo**, dei **conteggi** e della **deviazione standard**
- ❑ **Unione** dei due dataset rispetto all'Id del sensore
- ❑ **Verifica** delle medie ottenute
 - Rimuovo le rilevazioni con conteggio inferiore a 130, che corrispondono a meno di 21 ore di rilevazione in un giorno
- ❑ **Rimozione** delle righe corrispondenti a stazioni **non** ubicate **nella provincia di Milano**
- ❑ **Aggregazione** del dataset dei valori rispetto al singolo e al tipo di sensore in modo da unire le rilevazioni delle diverse stazioni
 - Calcolo della **media**, del **massimo** e della **deviazione standard**
- ❑ **Scelta della misurazione corretta** per i due tipi di sensore
 - **Valore massimo** giornaliero per la **temperatura**
 - **Valore medio** per la **radiazione globale**

Altri dataset utilizzati: descrizione e fonti

- ❑ [Dati sulle emissioni](#) mostra le **tonnellate emesse** per ogni inquinante, catalogate per le diverse **fonti antropiche** (rilevazioni compiute nel 2017)
 - Fonti divise per macrosettori, settori e attività
 - Dal sito INEMAR è possibile scaricare un dataset personalizzato, nel nostro caso relativo all'**Agglomerato di Milano** aggregando i dati per **macrosettori**
- ❑ [Dati pressione atmosferica](#) contiene i dati sulla **pressione atmosferica** dal 2016 al 2021 di una stazione ubicata a Milano in zona Zama
 - Dati aggregati **per giorno**

Scarica dati di emissione - Riassunto delle selezioni effettuate

Inquinanti selezionati <ul style="list-style-type: none">1 - Ossidi di zolfo (SO2 + SO3)2 - Ossidi di azoto (NO + NO2) come NO29 - Polveri con diametro <= 10 micron (PM10)33 - Polveri con diametro <= 2,5 micron (PM2.5)10429 - Totale precursori dell'ozono	Combustibili selezionati <ul style="list-style-type: none">11 - Altro12 - Senza combustibile1 - Benzina2 - Carbone3 - Diesel6 - GPL4 - Gas di raffineria5 - Gasolio7 - Kerosene8 - Legna e similari9 - Metano10 - Olio combustibile
Zone selezionate <ul style="list-style-type: none">1 - Agglomerato di Bergamo (Bergamo)4 - Zona A - pianura ad elevata urbanizzazione (Bergamo)5 - Zona B - pianura (Bergamo)6 - Zona C - montagna (Bergamo)7 - Zona D - fondovalle (Bergamo)2 - Agglomerato di Brescia (Brescia)4 - Zona A - pianura ad elevata urbanizzazione (Brescia)5 - Zona B - pianura (Brescia)6 - Zona C - montagna (Brescia)7 - Zona D - fondovalle (Brescia)3 - Agglomerato di Milano (Como)4 - Zona A - pianura ad elevata urbanizzazione (Como)6 - Zona C - montagna (Como)4 - Zona A - pianura ad elevata urbanizzazione (Cremona)5 - Zona B - pianura (Cremona)3 - Agglomerato di Milano (Lecco)4 - Zona A - pianura ad elevata urbanizzazione (Lecco)	Attività selezionate <ul style="list-style-type: none">1.1.1 - Caldaie con potenza termica >= 300 MW1.1.2 - Caldaie con potenza termica >= 50 e < 300 MW1.1.3 - Caldaie con potenza termica < 50 MW1.1.4 - Turbine a gas1.1.5 - Motori a combustione interna1.2.1 - Caldaie con potenza termica >= 300 MW1.2.2 - Caldaie con potenza termica >= 50 e < 300 MW1.2.3 - Caldaie con potenza termica < 50 MW1.2.4 - Turbine a gas1.2.5 - Motori a combustione interna1.3.1 - Caldaie con potenza termica >= 300 MW1.3.3 - Caldaie con potenza termica < 50 MW1.3.4 - Turbine a gas1.3.5 - Motori a combustione interna1.3.6 - Forni di raffineria1.5.4 - Turbine a gas2.1.3 - Caldaie con potenza termica < 50 MW

REALIZZAZIONE DELL'INFOGRAFICA: **DASHBOARD e VISUALIZZAZIONI**

Obiettivi dell'infografica

- ❑ Fornire le **informazioni** necessarie alla **comprensione globale** del **problema dell'inquinamento atmosferico** nella Val Padana e in particolare **nell'agglomerato urbano di Milano**
- ❑ Rappresentare **l'evoluzione della qualità dell'aria** (in termini di IQA) nell'agglomerato di Milano negli **ultimi 15 anni**, in **relazione** all'implementazione di **limiti legislativi** e norme, sempre più stringenti, in materia di emissioni
- ❑ Rappresentare gli **andamenti stagionali** della qualità dell'aria (in termini di IQA) e porla in **relazione** a fattori **meteorologici e antropici**
- ❑ Fornire una **visione di dettaglio** riguardo gli **inquinanti** atmosferici più **pericolosi** in termini di:
 - ❑ Caratteristiche e pericolosità
 - ❑ Principali fonti antropogeniche coinvolte
 - ❑ Andamento annuale e cause di eventuali innalzamenti o abbassamenti dei livelli rilevati
 - ❑ Andamento stagionale e cause meteorologiche e antropiche di eventuali innalzamenti o abbassamenti dei livelli rilevati

Strumenti utilizzati e struttura dell'infografica

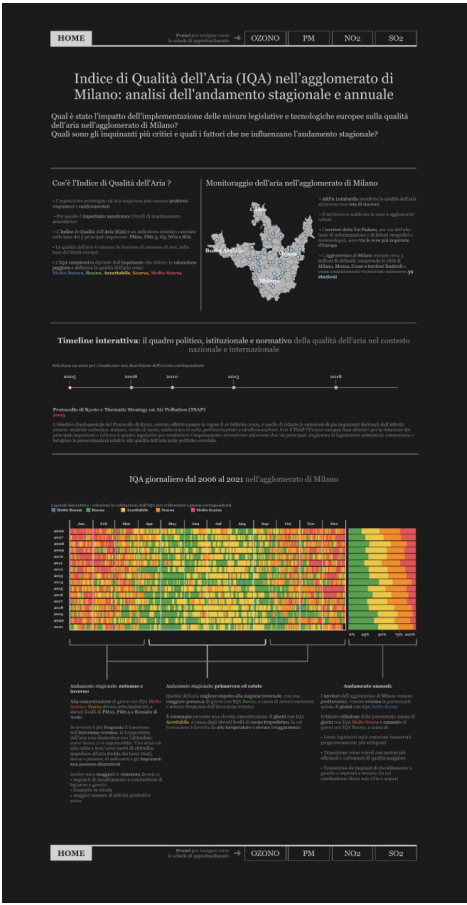
- ❑ Per lo sviluppo dell'infografica è stato utilizzato il software di analisi e visualizzazione **Tableau**
- ❑ L'infografica proposta è formata da **più visualizzazioni (worksheet) interattive** organizzate, insieme a **elementi testuali di context** di ausilio alla comprensione delle visualizzazioni, **in cinque pagine (dashboard)**
 - ❑ La prima pagina è fornisce una visione di insieme sul tema della qualità dell'aria e il suo andamento nell'agglomerato di milano, mentre le altre quattro pagine sono schede di approfondimento relative ai singoli inquinanti
 - ❑ Le pagine sono connesse tra loro mediante bottoni di navigazione integrati nelle pagine stesse:



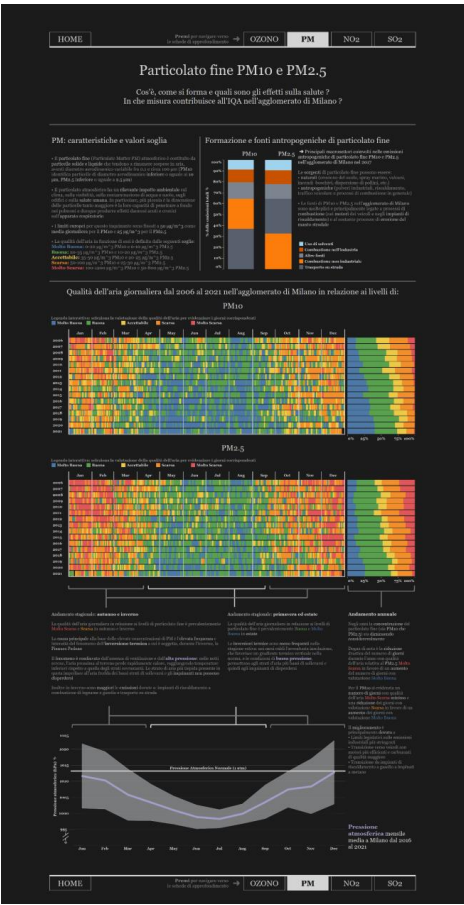
- ❑ L'infografica è stata pubblicata sulla piattaforma [Tableau Public](#)
- ❑ Di seguito verranno analizzate nel dettaglio le singole pagine dell'infografica, le sezioni in esse presenti e i grafici proposti

Le pagine dell'infografica

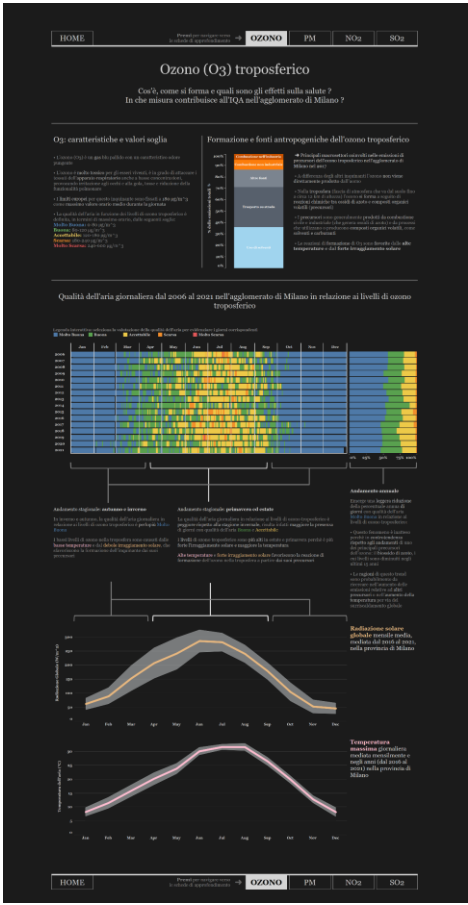
1/



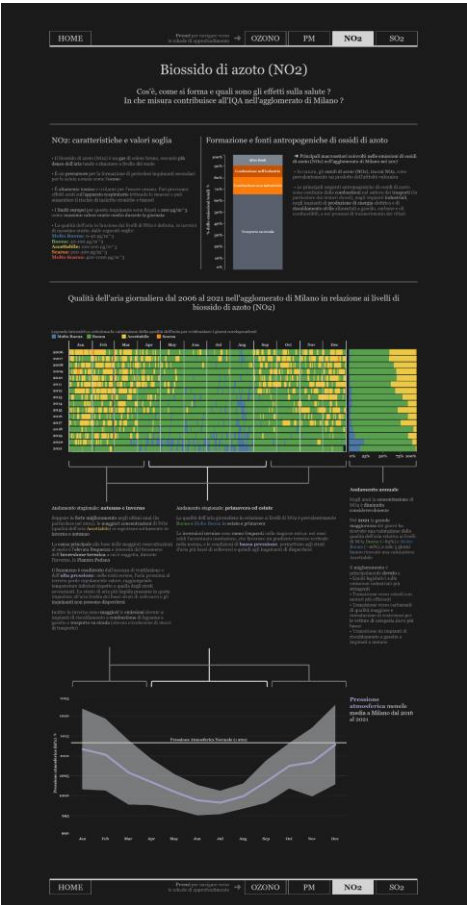
2/



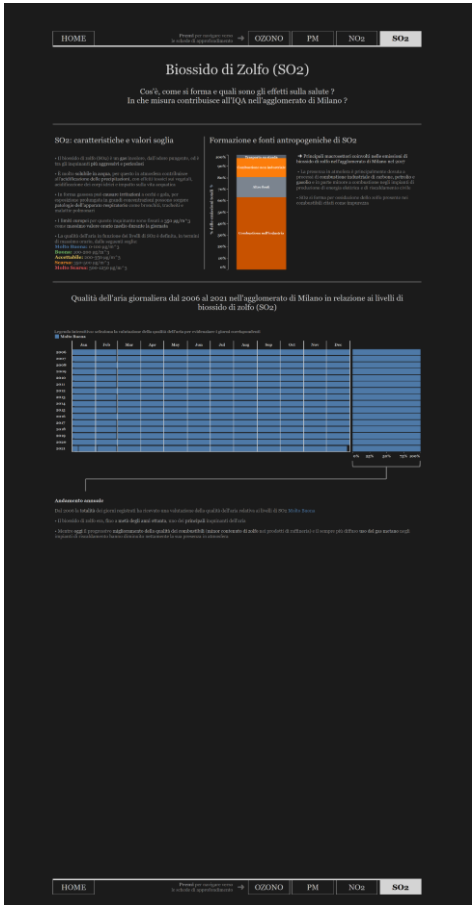
3/



4/

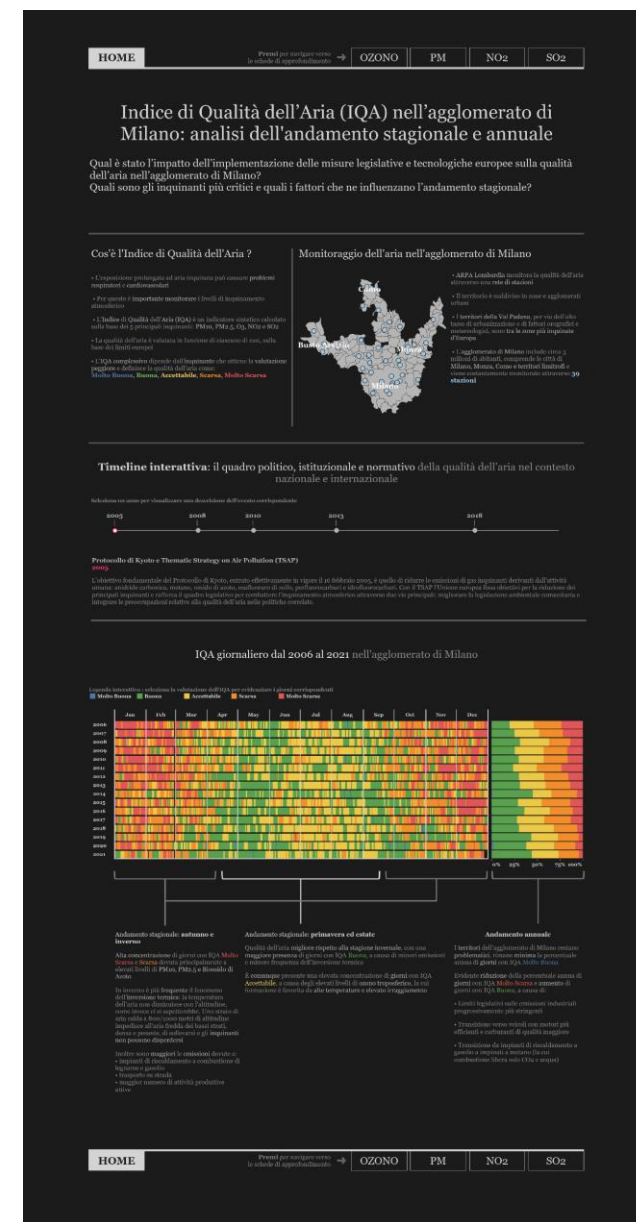


5/



Pagina iniziale: visione di insieme

- ❑ La prima pagina dell'infografica fornisce una visione di insieme riguardo il **tema dell'inquinamento atmosferico** e la sua declinazione specifica **nell'agglomerato di Milano**. La pagina presenta **diverse sezioni**:
1. In «**Cos'è l'Indice di Qualità dell'Aria ?**» viene descritto il significato e le modalità di calcolo dell'Indice di Qualità dell'Aria in Lombardia e gli inquinanti principali considerati
 2. In «**Monitoraggio dell'aria nell'agglomerato di Milano**» viene espresso sinteticamente il concetto di *zonizzazione* del territorio lombardo e vengono mostrati, con una mappa interattiva, i comuni inclusi nell'agglomerato urbano di Milano e le stazioni di monitoraggio ARPA
 3. Con la «**Timeline interattiva**: il quadro politico, istituzionale e normativo» vengono esposti i principali eventi e normative che hanno caratterizzato il panorama internazionale, europeo e nazionale in materia di inquinamento dell'aria
 4. Con la **visualizzazione «IQA giornaliero dal 2006 al 2021 nell'agglomerato di Milano»** viene mostrato l'andamento giornaliero e annuale dell'IQA nell'agglomerato di Milano negli ultimi 15 anni e con il *context* associato vengono esplorate le cause dei *trend* annuali e stagionali



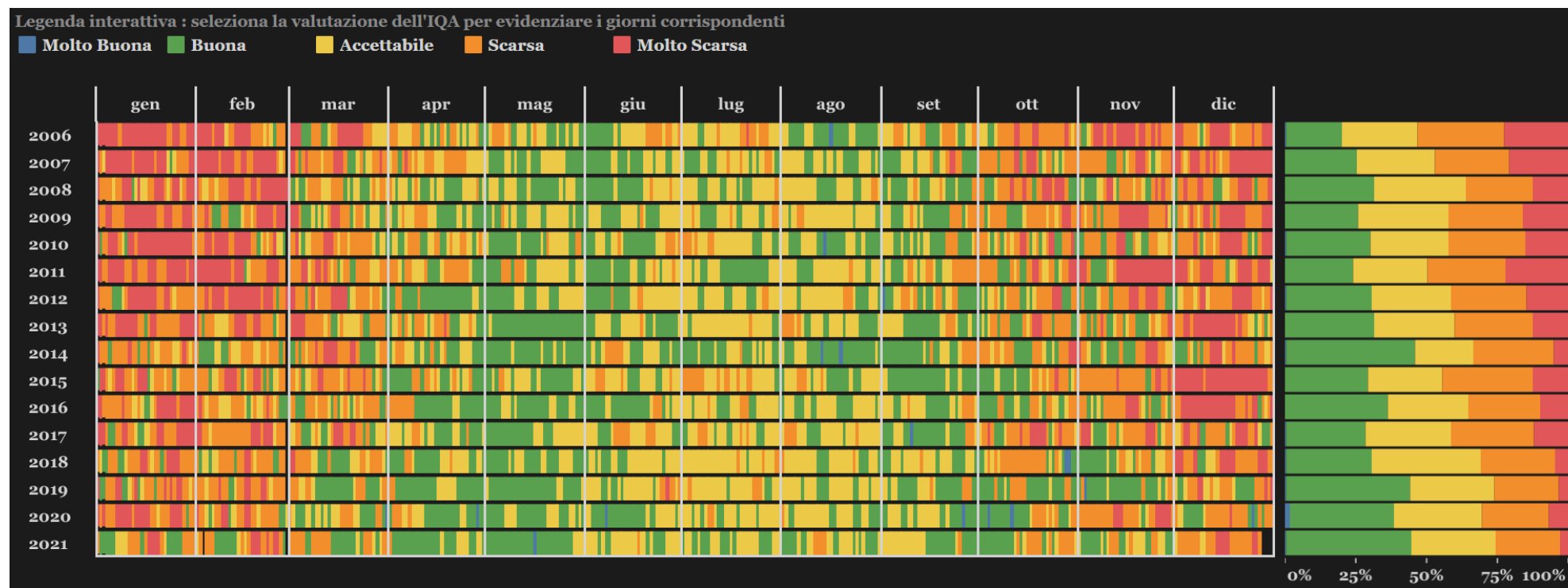
Schede di approfondimento inquinanti

- ❑ Le **altre quattro pagine** dell'infografica consistono in schede di approfondimento riguardanti i **singoli inquinanti** considerati nel calcolo dell'Indice di Qualità dell'Aria: particolato fine (PM10 e PM2.5), ozono troposferico (O3), biossido di azoto (NO2) e biossido di zolfo (SO2)
- ❑ Tutte le schede di approfondimento presentano la **medesima struttura**:
 1. Nella sezione «**caratteristiche e valori soglia**» vengono descritte le proprietà organolettiche, gli **effetti sulla salute umana** e sull'ambiente dei singoli inquinanti. Inoltre vengono esposti i limiti e le soglie europee che permettono di definire la qualità dell'aria durante una giornata come **Molto Buona**, **Buona**, **Accettabile**, **Scarsa**, **Molto Scarsa** in funzione del singolo inquinante
 2. In «**Formazione e fonti antropogeniche**» sono descritti i principali **processi di formazione** dell'inquinante in atmosfera e **visualizzati i macrosettori coinvolti** nelle emissioni dell'inquinante nell'agglomerato urbano di Milano
 3. Con la **visualizzazione «Qualità dell'aria giornaliera** in relazione ai livelli del singolo inquinante» viene mostrato l'andamento giornaliero e annuale dell'IQA nell'agglomerato di Milano negli ultimi 15 anni e con il *context* associato vengono esplorate le cause dei *trend* annuali e stagionali
 4. Il context relativo all'andamento stagionale dei livelli di inquinante viene arricchito con **visualizzazioni** che descrivono l'**andamento stagionale delle proprietà meteorologiche** associate a significative variazioni nella concentrazione dell'inquinante



Visualizzazione principale: obiettivo

- ❑ La visualizzazione principale ha l'obiettivo di **mostrare l'andamento giornaliero e annuale della qualità dell'aria nell'agglomerato urbano di Milano dal 2006 a oggi**
- ❑ La visualizzazione è utilizzata **sia** per mostrare la qualità dell'aria complessiva in termini di **IQA** nella pagina principale, **che** per visualizzare la qualità dell'aria in relazione al **singolo inquinante** nelle schede di approfondimento

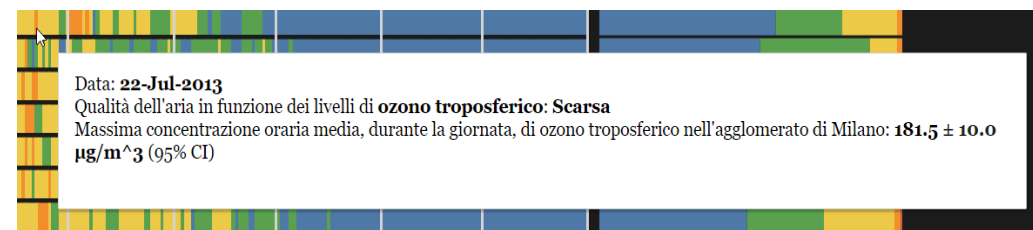


Visualizzazione principale: caratteristiche

- ❑ La visualizzazione principale è **composta da due visualizzazioni** affiancate associate alla trasmissione di diverse informazioni: una **heat map** e uno **100% stacked bar chart**
- ❑ Nella heat map viene associato ad **ogni giorno dell'anno** (ogni cella della matrice) un **colore** corrispondente a una delle **modalità della scala** di valutazione della **qualità dell'aria**
 - ❑ È stata scelta una **palette di colori discreta** che permettesse di **esprimere** la **natura ordinale** della **variabile categoriale** della qualità dell'aria e che permettesse di associare, visivamente, a una qualità dell'aria "Scarsa" e "Molto Scarsa" e "Accettabile" un'accezione negativa e una positiva a "Molto Buona" e "Buona"
 - ❑ La visualizzazione consente di osservare la distribuzione delle valutazioni di qualità dell'aria nei diversi mesi e nelle diverse stagioni, permettendo all'osservatore di **individuare eventuali concentrazioni di giorni con qualità dell'aria simile**
 - ❑ La sovrapposizione delle diverse righe della matrice della heat map, relative ai diversi anni studiati, permette di evidenziare la presenza di **pattern stagionali ricorrenti** ed eventuali miglioramenti/peggioramenti in certi periodi dell'anno
- ❑ Lo scopo del **100% stacked bar chart** è quello di mostrare un'informazione non facilmente deducibile dalla heat map, ovvero la **percentuale di giorni** classificati con una certa valutazione di qualità dell'aria **sul totale** di giorni registrati durante l'anno
 - ❑ Le diverse modalità della scala di qualità dell'aria sono state ordinate
 - ❑ La sovrapposizione delle barre relative ad anni successivi permette di evidenziare un eventuale miglioramento/peggioramento della qualità dell'aria negli anni

Visualizzazione principale: interazioni

- ❑ Posizionando il cursore su una delle celle della heat map (nella visualizzazione relativa all'**IQA** nella pagina principale) viene attivata l'apertura di un **tooltip** contenente informazioni di dettaglio su:
 - ❑ **IQA** nel giorno specifico
 - ❑ **Qualità dell'aria** relativa ai **singoli inquinanti**: consentendo all'utente di determinare il contributo dei diversi inquinanti alla qualità dell'aria complessiva
- ❑ Posizionando il cursore su una delle celle della heat map (nelle pagine relative ai **singoli inquinanti**) viene attivata l'apertura di un **tooltip** contenente informazioni di dettaglio su:
 - ❑ Qualità dell'aria relativa al singolo inquinante nel giorno specifico
 - ❑ **Concentrazione** di inquinante nel giorno specifico, associata a un valore di **incertezza** nella stima (intervallo di confidenza al 95%)

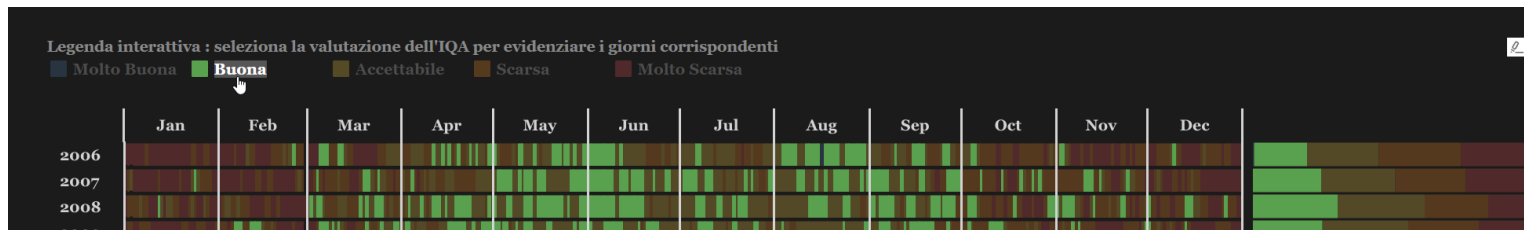


Visualizzazione principale: interazioni

- ❑ Posizionando il cursore su una delle sezioni di ogni barra del **100% stacked bar chart** (nella visualizzazione relativa all'IQA nella pagina principale e anche nelle schede di approfondimento) viene attivata l'apertura di un **tooltip** contenente informazioni di dettaglio riguardo il **numero e la percentuale di giorni** classificati con una certa valutazione di qualità dell'aria sul totale di giorni registrati durante l'anno

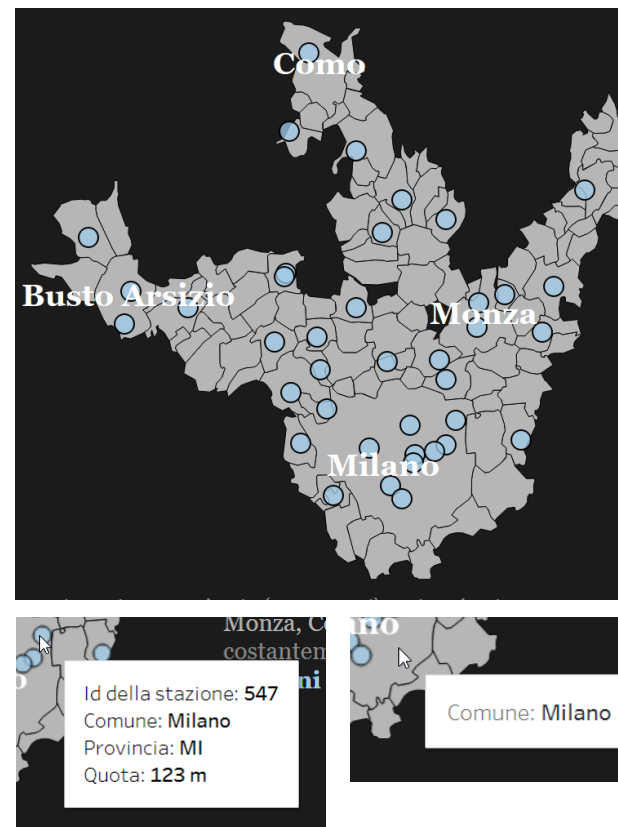


- ☐ Premendo sugli **elementi della legenda** associate alle diverse valutazioni della qualità dell'aria è possibile **evidenziare** le porzioni di grafico corrispondenti (sia nel bar chart che nella heat map)



Visualizzazioni secondarie: mappa agglomerato di Milano

- ❑ Nella pagina principale è presente una mappa **interattiva dei comuni** dell'agglomerato di Milano e delle **stazioni di monitoraggio ARPA** presenti sul territorio
- ❑ Passando con il cursore sul territorio dei comuni viene mostrato un tooltip contenente il nome del comune
- ❑ Passando con il cursore sulle singole stazioni (indicate con cerchi di colore azzurro) viene mostrato un tooltip con alcuni dettagli sulle stesse



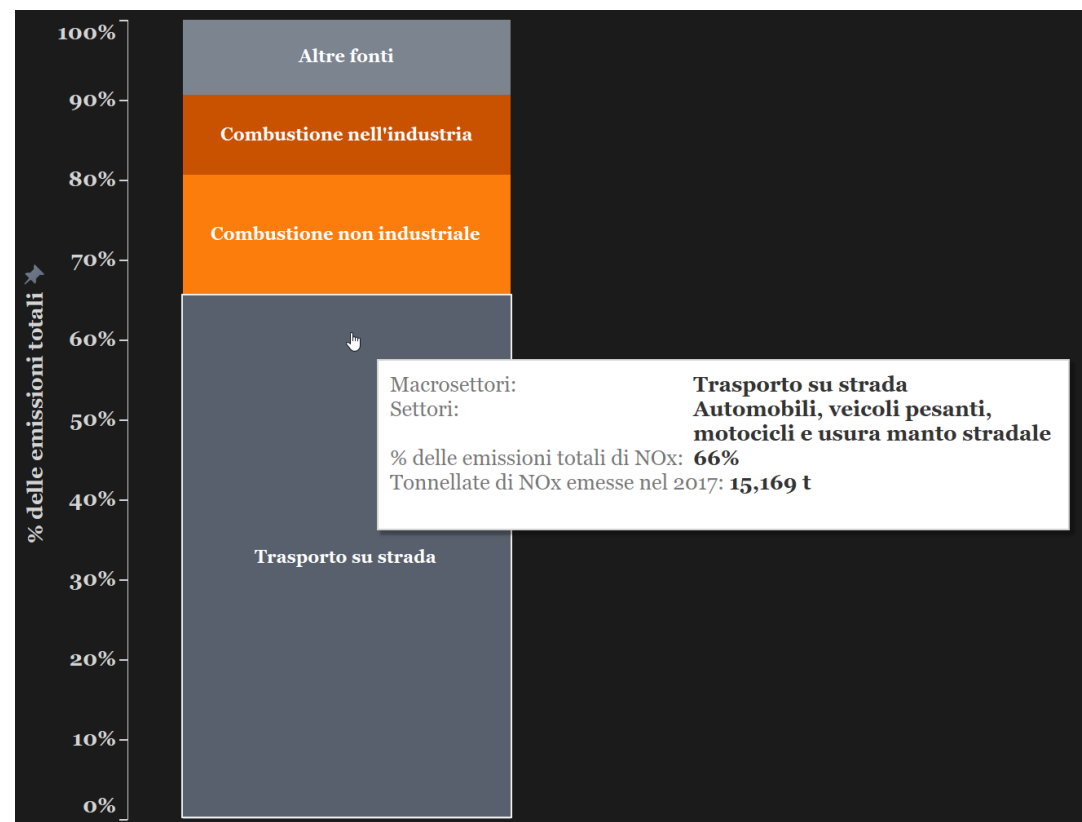
Visualizzazioni secondarie: timeline interattiva

- ❑ Nella sezione «Timeline interattiva: il quadro politico, istituzionale e normativo» della pagina principale é stata inserita una timeline interattiva che prevede **cinque anni selezionabili**
- ❑ La visualizzazione ha l'obiettivo di **esporre i principali eventi e normative** che hanno caratterizzato il panorama internazionale, europeo e nazionale in materia di **inquinamento atmosferico** e il suo impatto sulla **salute umana**
- ❑ **Interazione:**
 - ❑ premendo sugli elementi della timeline è possibile visualizzare un titolo e una breve descrizione relativa all'evento (o gli eventi) corrispondente
 - ❑ La porzione di timeline relativa al 'passato' (rispetto all'evento selezionato) assume un colore acceso, la porzione relativa al 'futuro' mantiene una colorazione scura



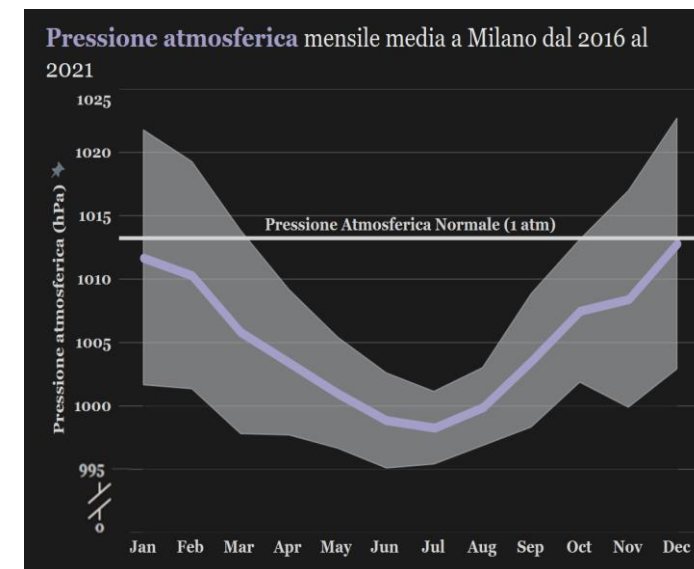
Visualizzazioni secondarie: fonti antropogeniche delle emissioni dei principali inquinanti

- ❑ In ogni **scheda di approfondimento** relativa a un singolo inquinante, nella sezione «Formazione e fonti antropogeniche», è stato inserito un **100% stacked bar chart**
- ❑ La visualizzazione ha l'obiettivo di mostrare i **principali macrosettori** (industriali e non) coinvolti nelle **emissioni** dello specifico inquinante e quindi fornire informazioni utili alla comprensione degli andamenti dei livelli di inquinanti descritti nella visualizzazione principale
- ❑ **Interazione:** il passaggio del cursore sulle diverse porzioni di ogni barra provoca l'apertura di un **tooltip** contenente informazioni di dettaglio riguardo:
 - ❑ I fenomeni e i settori specifici che caratterizzano ogni macrosettore
 - ❑ La percentuale di emissioni attribuibile al macrosettore rispetto al totale
 - ❑ La massa di inquinante emessa, attribuibile al macrosettore nell'anno di riferimento



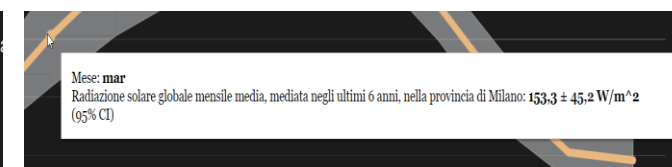
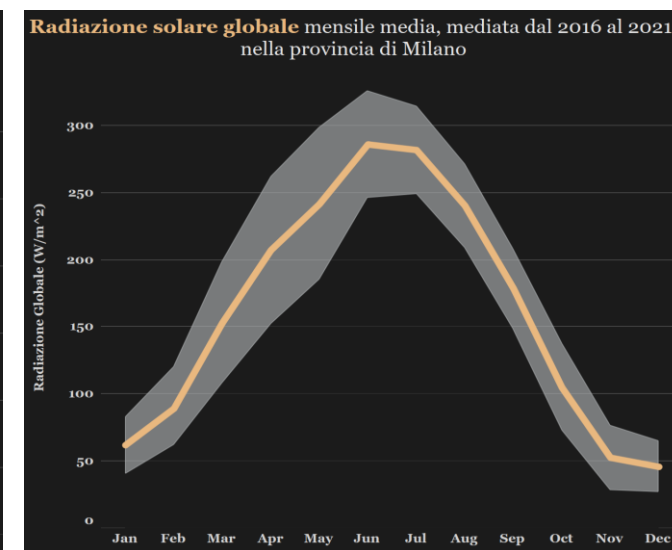
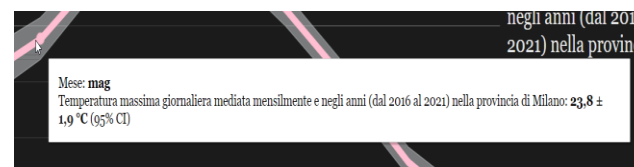
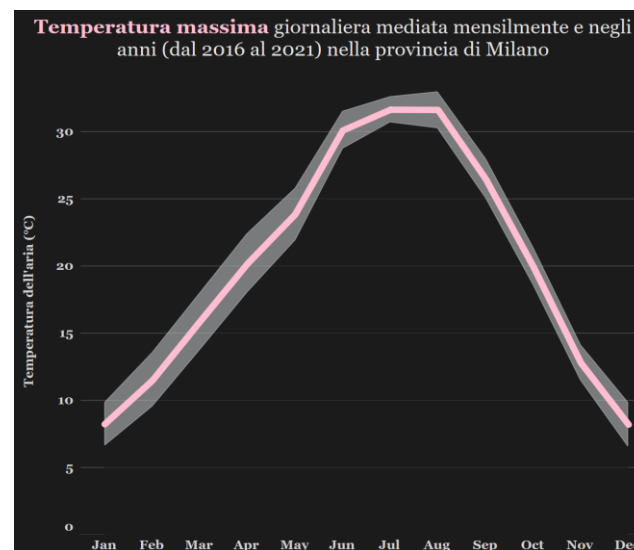
Visualizzazioni secondarie: andamenti delle variabili meteorologiche

- ❑ Nelle **schede di approfondimento** dei singoli inquinanti, il *context* relativo agli andamenti stagionali è stato integrato con alcune visualizzazioni riguardanti gli andamenti delle principali variabili meteorologiche che possono determinare variazioni significative nella concentrazione dell'inquinante
- ❑ La variabile studiata in relazione all'**andamento stagionale** delle concentrazioni di **NO2, PM10 e PM2.5** è la **pressione atmosferica**
 - ❑ Le condizioni di **alta pressione** e l'assenza di ventilazione, che caratterizzano la stagione invernale nella Pianura Padana, causano una maggiore frequenza e intensità del fenomeno dell'**inversione termica**, che **impedisce** una normale **dispersione** degli **inquinanti** considerati
 - ❑ Si è scelto quindi di usare un **line chart** per rappresentare l'andamento della pressione atmosferica mensile media, mediata ulteriormente negli ultimi 5 anni, a Milano
 - ❑ L'asse delle **ordinate** è stato **troncato**. È stata però scelta una **linea di riferimento** (più utile rispetto alla pressione atmosferica nulla) in corrispondenza del valore di pressione atmosferica normale (1 atm) al livello del mare: al di sotto di essa sono presenti stati di bassa pressione, al di sopra di alta pressione
 - ❑ È stata rappresentata la **dispersione dei dati intorno alla media** in termini di **deviazione standard**: lo scopo è quello di mostrare come fenomeni di alta pressione siano più frequenti in inverno piuttosto che in estate



Visualizzazioni secondarie: andamenti delle variabili meteorologiche

- ❑ Le variabili studiate, in relazione all'andamento stagionale delle concentrazioni di **O3**, sono la **temperatura massima giornaliera** e la **radiazione solare globale** nella provincia di Milano
- ❑ Alte temperature e forte irraggiamento solare favoriscono infatti la reazione di formazione dell'ozono nella troposfera a partire dai suoi precursori
- ❑ Le misurazioni di entrambe le variabili sono state **medie mensilmente e negli ultimi 5 anni**
- ❑ Per la visualizzazione si è scelto di usare un **line chart**
- ❑ Il valore medio è stato accompagnato da un valore di incertezza nella stima (intervallo di confidenza al 95%) rappresentato visivamente attraverso una banda e specificato, insieme al valore medio in un tooltip



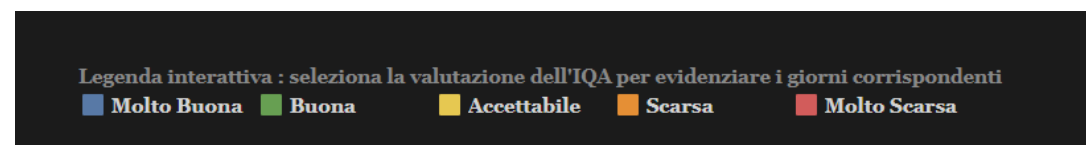
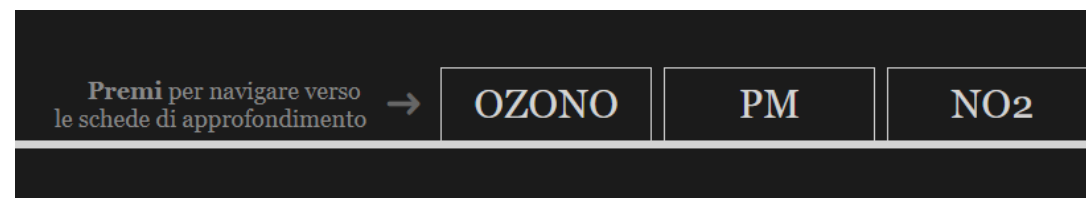
VALUTAZIONE DI QUALITÀ

Valutazione euristica

- ❑ **Problemi di usabilità emersi**, ed in seguito risolti, coinvolgendo **6 utenti** esterni al progetto:
 - **Difficoltà nel riconoscere** la presenza di tasti ed **elementi interattivi**
 - **Colorazione** controintuitiva della **timeline**
 - **Basso contrasto tra le diverse linee di riferimento** al context nella visualizzazione principale
 - **Mancanza riferimento cromatico** nei grafici di **temperatura e radiazione solare**

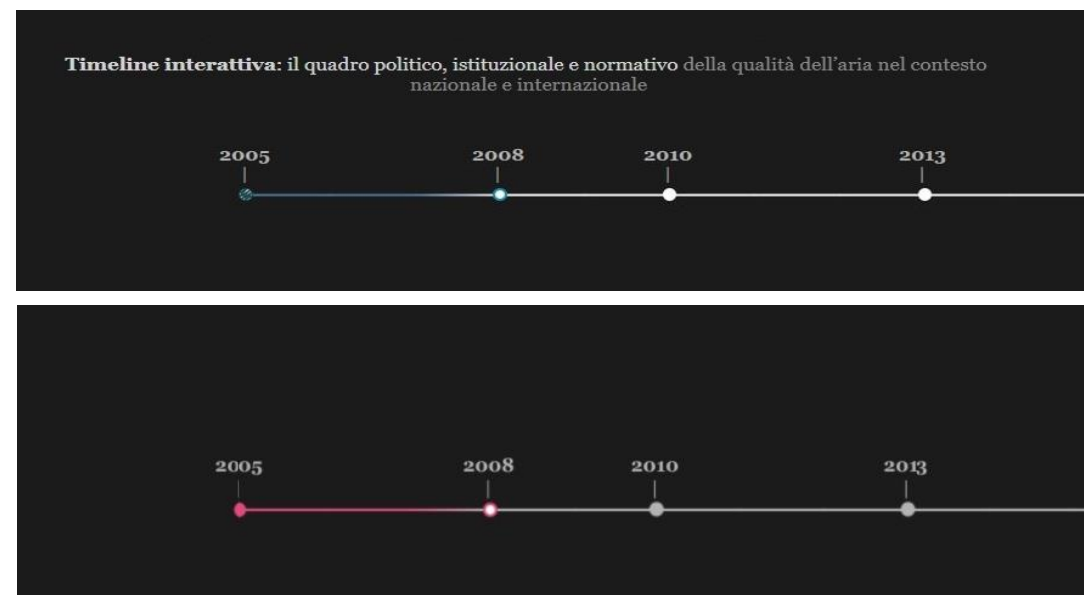
Valutazione euristica: etichette elementi interattivi

- ❑ Sono state riscontrate delle difficoltà nel riconoscere quali fossero gli **elementi interattivi** della visualizzazione poiché **non erano adeguatamente segnalati**
- ❑ Questa problematica è stata risolta con l'**aggiunta di etichette** in prossimità degli elementi interattivi in modo da aumentare la fruibilità



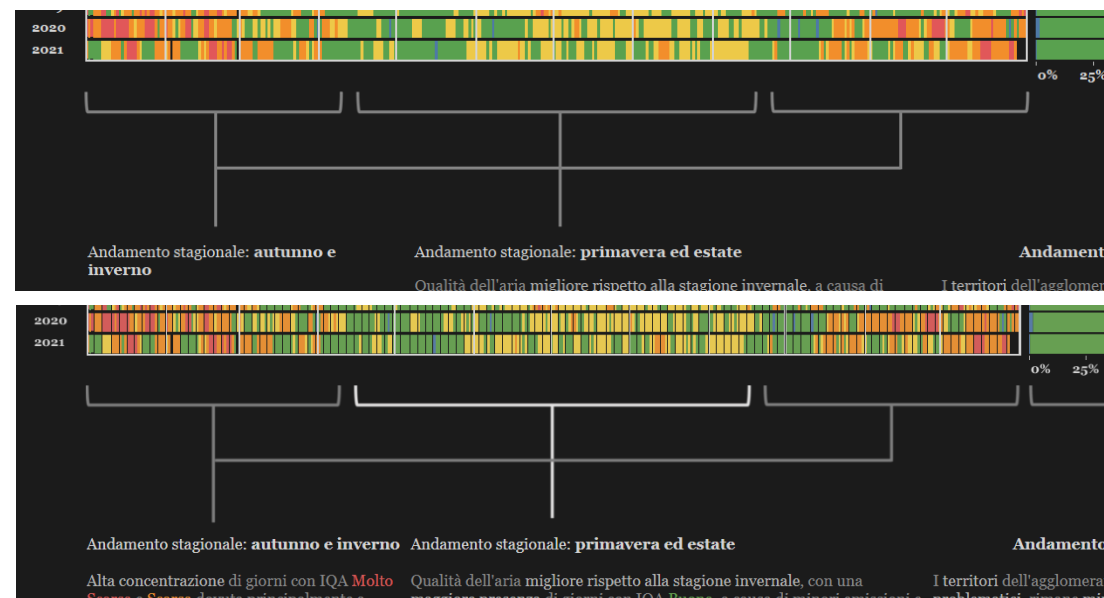
Valutazione euristica: colorazione timeline

- ❑ La **colorazione originale** della timeline prevedeva un colore blu scuro che, venendo confuso per grigio, sembrava definire un andamento a ritroso nel tempo e quindi **controintuitivo**
- ❑ Questa problematica è stata **risolta cambiando il colore** del periodo 'passato' rispetto all'evento selezionato



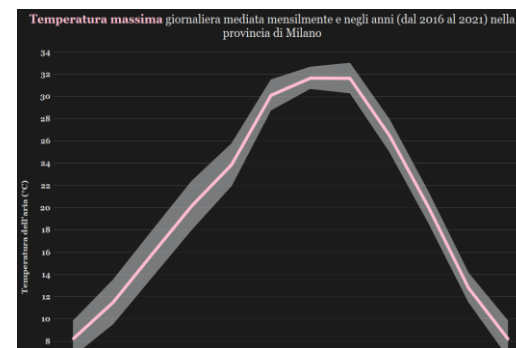
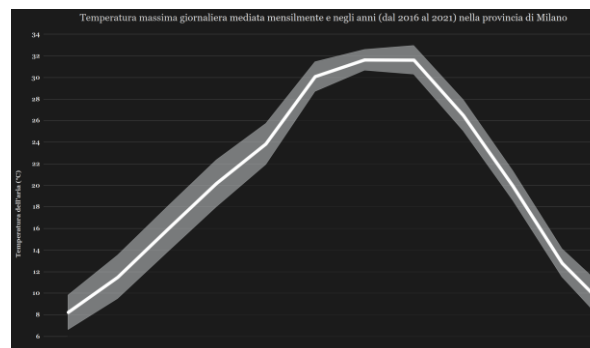
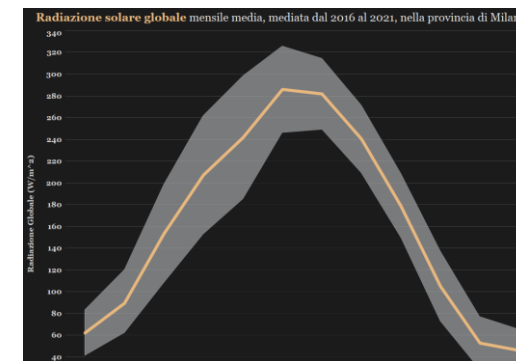
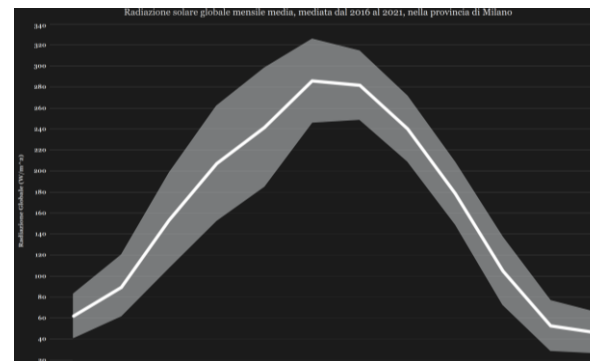
Valutazione euristica: riferimenti ai mesi

- ❑ Inizialmente i **riferimenti al context** mensile della visualizzazione principale erano dello **stesso colore** e non permettevano a colpo d'occhio di rintracciare la didascalia corretta
- ❑ Questa problematica è stata **risolta differenziando i colori** dei riferimenti



Valutazione euristica: mancanza riferimento cromatico

- ❑ I **line chart** interessati da questo problema sono quelli nelle pagine relative all'ozono troposferico, al PM e al biossido di azoto
- ❑ In queste pagine le visualizzazioni relative alle variabili meteorologiche di **temperatura** massima e **radiazione** solare globale sono **sovrapposti**
- ❑ Anche in questo caso si tratta di una **cattiva scelta dei colori** che non permette di riconoscere a colpo d'occhio il grafico a cui si fa riferimento nel context: inizialmente le linee presenti nei due grafici erano rappresentate con lo stesso colore
- ❑ Questa problematica è stata **risolta differenziando i colori** delle linee nei due grafici per poterli richiamare più facilmente all'interno del context (colorando il testo con il colore corrispondente)

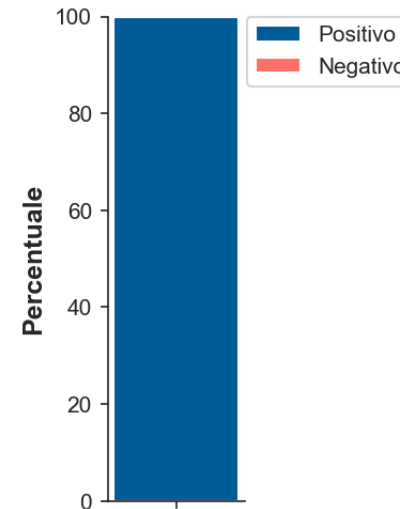


User test: Task 1

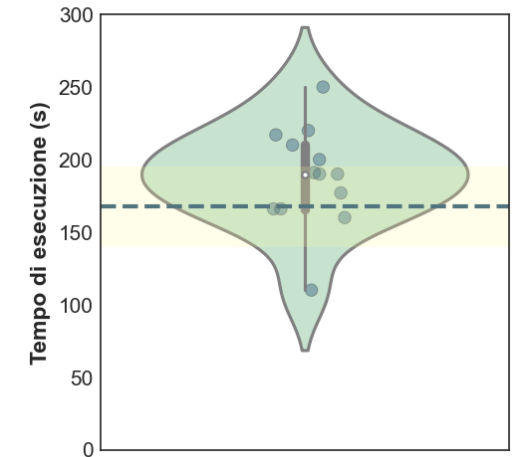
Mira a valutare la corretta interpretazione della visualizzazione principale, in particolare sfruttando l'interazione della bar chart che mostra **la percentuale di giorni** classificati con una certa valutazione in un certo anno

- ❑ «Dal 2006 ad 2021 quale anno ha avuto il maggior numero di giorni con una qualità dell'aria classificata 'molto scarsa'?»
 - **Tempo di esecuzione ottimale** ottenuto mediando i tempi dei 4 sviluppatori: $167.7 \pm 19.5s$ (IC 95%)
 - Valore medio del **tempo impiegato dai 13 utenti testati**: $188.3 \pm 18.0s$ (IC 95%)
- ❑ Il task è stato **portato a termine da tutti gli utenti** in un tempo simile a quello ottimale

Esiti primo task



Distribuzione dei tempi di risposta rispetto al "valore ottimale" del primo task (13 utenti)

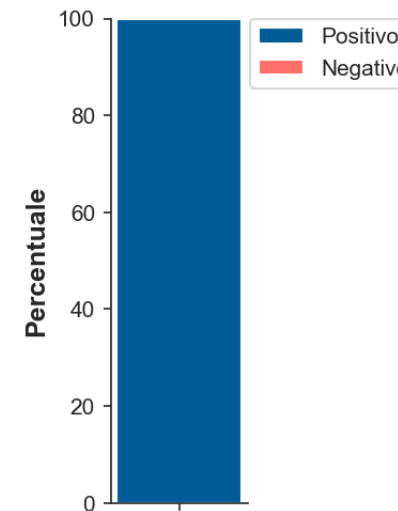


User test: Task 2

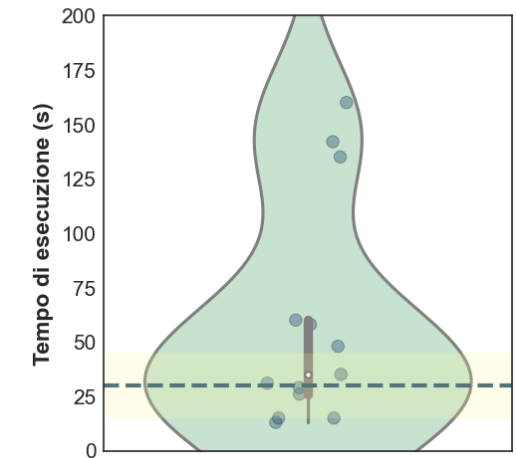
Mira a valutare la corretta interpretazione della visualizzazione principale (in particolare della heatmap), per quanto riguarda gli andamenti stagionali della qualità dell'aria

- ❑ «*Quale inquinante influenza maggiormente il punteggio nei mesi estivi? E quale fenomeno meteorologico è fortemente correlato alla sua concentrazione?*»
 - **Tempo di esecuzione ottimale** ottenuto mediando i tempi dei 4 sviluppatori: $29.9 \pm 15.7s$ (IC 95%)
 - Valore medio del **tempo impiegato dai 13 utenti testati**: $59.0 \pm 27.1s$ (IC 95%)
- ❑ Il task è stato **portato a termine da tutti gli utenti**, da notare come 3 utenti abbiano impiegato un tempo molto elevato poiché hanno navigato le schede relative ai singoli inquinanti per dare la risposta e non hanno utilizzato l'interazione e il context della visualizzazione nella pagina principale

Esiti Secondo task



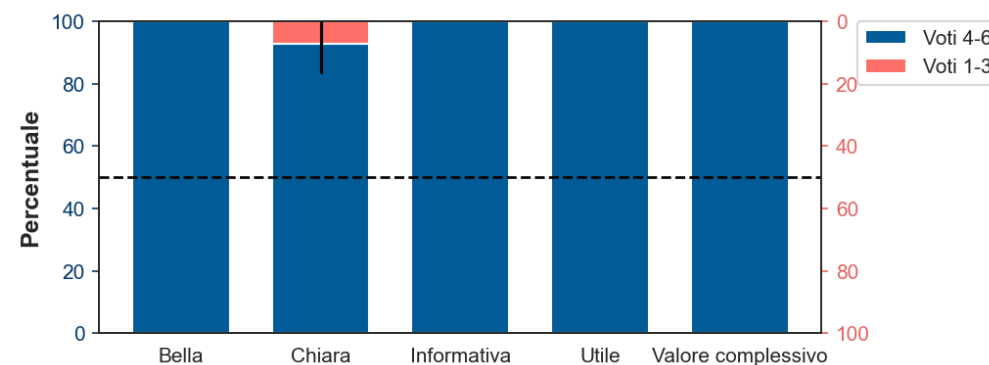
Distribuzione dei tempi di risposta rispetto al "valore ottimale" del secondo task (13 utenti)



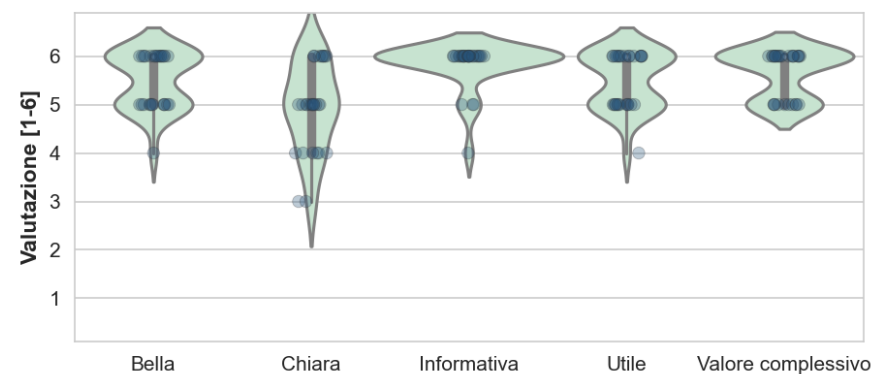
Questionario psicometrico: distribuzione dei voti

- ❑ Il questionario psicometrico è stato somministrato a **28 utenti** e mira a valutare la qualità della visualizzazione
- ❑ È basato sulla **scala Cabitza-Locoro** che mira a valutare utilità, chiarezza, informatività e bellezza della visualizzazione
 - Le **votazioni** rispetto ai singoli aspetti sono state **tutte positive** (punteggio >3) **tranne** per l'aspetto della **chiarezza** che ha ricevuto due voti discreti
 - Dai *violin plot* si può apprezzare la **distribuzione dei voti** nei diversi aspetti valutati: tutti i parametri risultano **molto alti** (5 e 6) mentre come già notato l'aspetto della **chiarezza** è stato **leggermente penalizzato**

Voti positivi vs voti negativi del sondaggio (28 questionari)



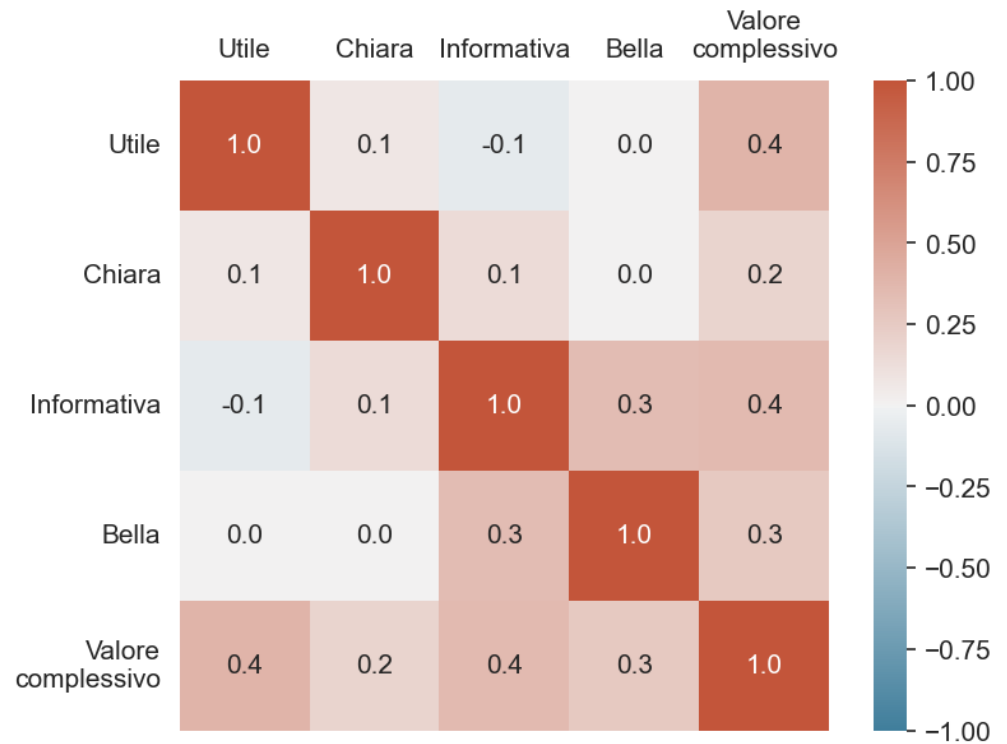
Distribuzione dei voti del questionario per ogni aggettivo (28 questionari)



Questionario psicometrico: correlazione tra i voti

- ❑ Dal grafico delle **correlazioni** tra i voti si nota che i gli attributi più legati **al valore complessivo dell'infografica** sono quelli dell'**utilità** e dell'**informatività**
- ❑ Data la **correlazione positiva** si evince quindi come chi ha valutato positivamente la visualizzazione abbia anche apprezzato l'utilità e l'informatività della stessa

Correlazione tra i voti del questionario
(28 questionari complessivi)



Bibliografia e sitografia

- ❑ Actis-Grosso R. & Batini C. Locoro A. Cabitza F. "Static and interactive infographics in daily tasks: A value-in-use and quality of interaction user study." In: Computers in Human Behavior (2017).
- ❑ ARPA Lombardia. (2021). LA VALUTAZIONE DELLA QUALITÀ DELL'ARIA IN LOMBARDIA.
<https://www.arpalombardia.it/Pages/Aria/Da-sapere.aspx>
- ❑ Dlgs 155/2010 - Attuazione della direttiva 2008/50/CE relativa alla qualità dell'aria ambiente e per un'aria più pulita in Europa (15 settembre 2010) <https://web.camera.it/parlam/leggi/deleghe/testi/10155dl.htm>
- ❑ European Commission. (2021). Existing Legislation - Air Quality - Environment - European Commission. Ec.Europa.Eu.
https://ec.europa.eu/environment/air/quality/existing_leg.htm
- ❑ Indice di qualità dell'aria | ARPA Lombardia <https://www.arpalombardia.it/Pages/Aria/Modellistica/Indice-qualit%C3%A0-aria.aspx>

Bibliografia

- ❑ Zonizzazione | ARPA Lombardia <https://www.arpalombardia.it/Pages/Aria/Rete-di-rilevamento/Zonizzazione.aspx?firstlevel=Rete%20di%20rilevamento>
- ❑ DIRETTIVA (UE) 2016/2284 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO. (2016). Gazzetta Ufficiale Dell'Unione Europea. <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32016L2284&from=IT>
- ❑ Commission to the Council and the European Parliament. (2005). Thematic Strategy on Air Pollution. D. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:l28159>
- ❑ GARD ITALIA – INQUINAMENTO ATMOSFERICO E CAMBIAMENTI CLIMATICI: Elementi per una strategia nazionale di prevenzione. https://www.salute.gov.it/imgs/C_17_pubblicazioni_2945_allegato.pdf