

# Progettazione di Database sugli Aeroporti

Enrico Guerriero, EC2100780, Statistica e Informatica per l'Azienda, la Finanza e l'Assicurazione  
Appello del 12 luglio 2023

## 1 Introduzione

L'obiettivo di questo progetto è quello di creare un database che fornisca tutte le informazioni che possono risultare utili riguardo i dati di volo.

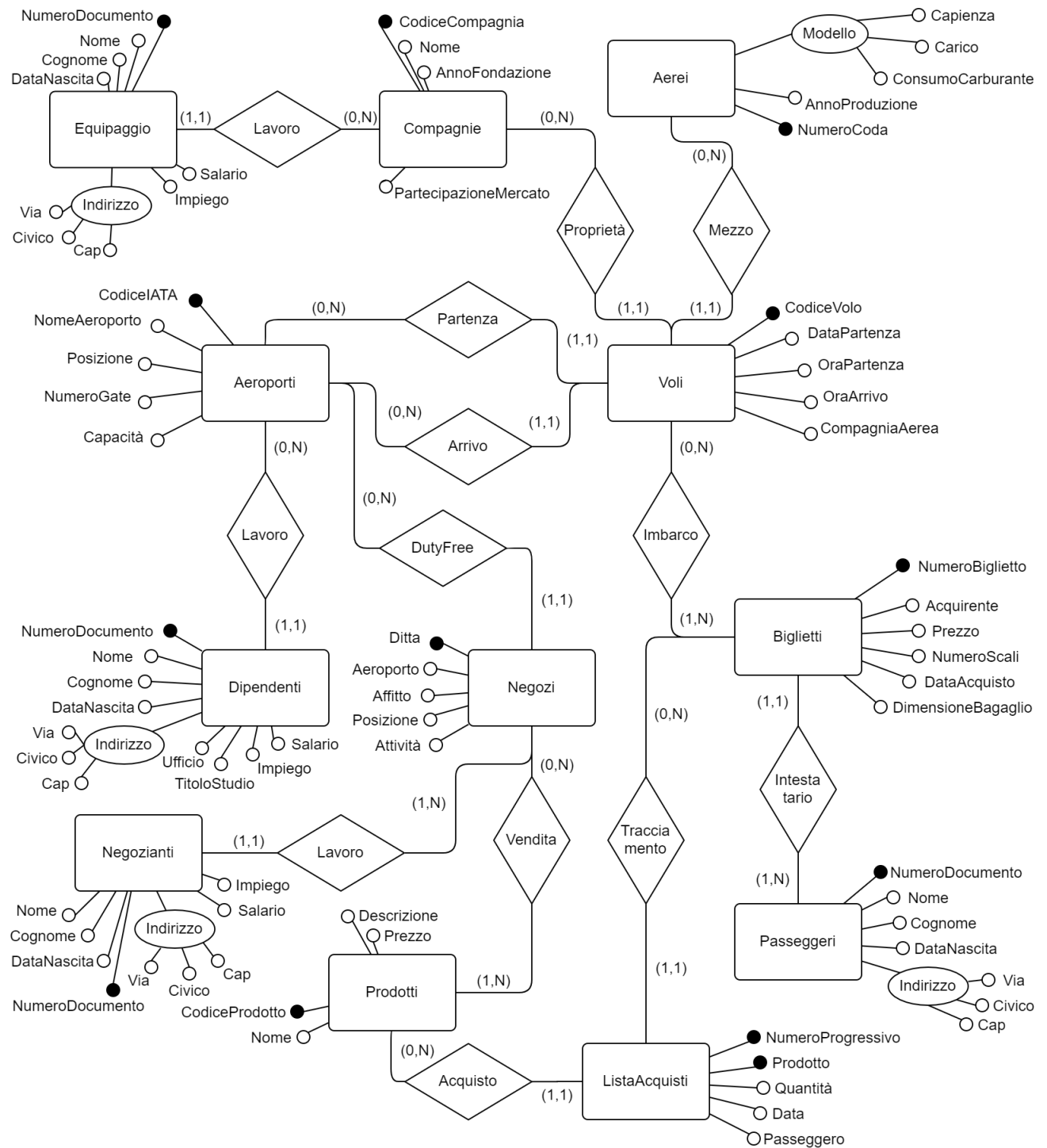
Per semplicità, si ipotizza di avere un database con informazioni su tutti gli aeroporti italiani e tutti i voli interni al territorio italiano (quindi con partenza e arrivo in aeroporti italiani).

Il database è pensato ad un'estensione internazionale, per esempio europea, quindi sono presenti variabili, come il numero di scali, che hanno davvero senso solo su territori più estesi; anche i codici identificativi utilizzati nel database sono internazionali.

### Requisiti e ipotesi

- Ciascun aeroporto è dotato di codice IATA, identificativo dell'aeroporto; dell'aeroporto si conosce il nome, la posizione geografica (regione di appartenenza), il numero di gate presenti e la capacità in termini di persone.
- In ogni aeroporto sono previsti dei voli, identificati dal codice di volo, in partenza o in arrivo; dei voli si conosce la data, l'ora di partenza, l'ora di arrivo, l'aereo che ha effettuato il volo e la compagnia aerea.
- I voli sono effettuati dagli aerei, identificati dal numero di coda, dei quali si vuole conoscere anno di produzione e tre dati legati al modello dell'aereo: la capienza, il carico e il consumo di carburante.
- La compagnia aerea che effettua i voli è identificata da un codice (anche in questo caso un codice IATA, ma diverso da quello degli aeroporti) e si vuole conoscere il nome della compagnia, l'anno di fondazione e la quota di partecipazione di mercato.
- Le compagnie aeree hanno dei dipendenti, identificati dal numero del documento di identità, dei quali si vuole conoscere nome, cognome, data di nascita, indirizzo, impiego e stipendio.
- Per ogni volo sono necessari dei biglietti, identificati da un numero di biglietto e dal volo, dei quali si vuole conoscere la data di acquisto, l'acquirente, il prezzo, la dimensione del bagaglio prevista dal biglietto e il numero di scali previsti.
- I passeggeri sono identificati dal numero di documento di identità, e si vuole conoscere nome, cognome, data di nascita e indirizzo.
- Nell'aeroporto sono presenti dei dipendenti, identificati dal numero di documento di identità, e si vuole conoscere nome, cognome, data di nascita, indirizzo, titolo di studio, impiego, stipendio e, se hanno un ufficio, dove si trova.
- Nell'aeroporto sono presenti delle attività commerciali, all'interno del duty-free, identificate dal nome della ditta, e delle quali si conosce il tipo di attività, la posizione nell'aeroporto, la durata e l'inizio del contratto d'affitto.
- Nelle attività commerciali sono presenti dei lavoratori, identificati dal numero del documento di identità, dei quali si vuole conoscere nome, cognome, indirizzo, data di nascita, ditta e aeroporto per i quali lavorano, impiego e stipendio.
- Nelle attività vengono venduti dei prodotti, identificati da un codice, dei quali si vuole conoscere nome, ditta che li produce, la descrizione del prodotto e il prezzo.
- Per ciascun acquisto di ciascun prodotto è presente una lista acquisti, identificata dal numero progressivo dello scontrino e il prodotto venduto, e si conosce la quantità di prodotto venduta, la data di vendita e il passeggero che ha effettuato l'acquisto; quest'ultima è un'ipotesi fatta da me, è un'ipotesi molto forte ma coerente, poiché vengono considerate solo le attività commerciali interne al duty-free e non quelle di tutto l'aeroporto, quindi si tratta di attività cui hanno accesso soltanto passeggeri muniti di biglietto.

## 2 Primo schema ER



Schema Entity Relationship

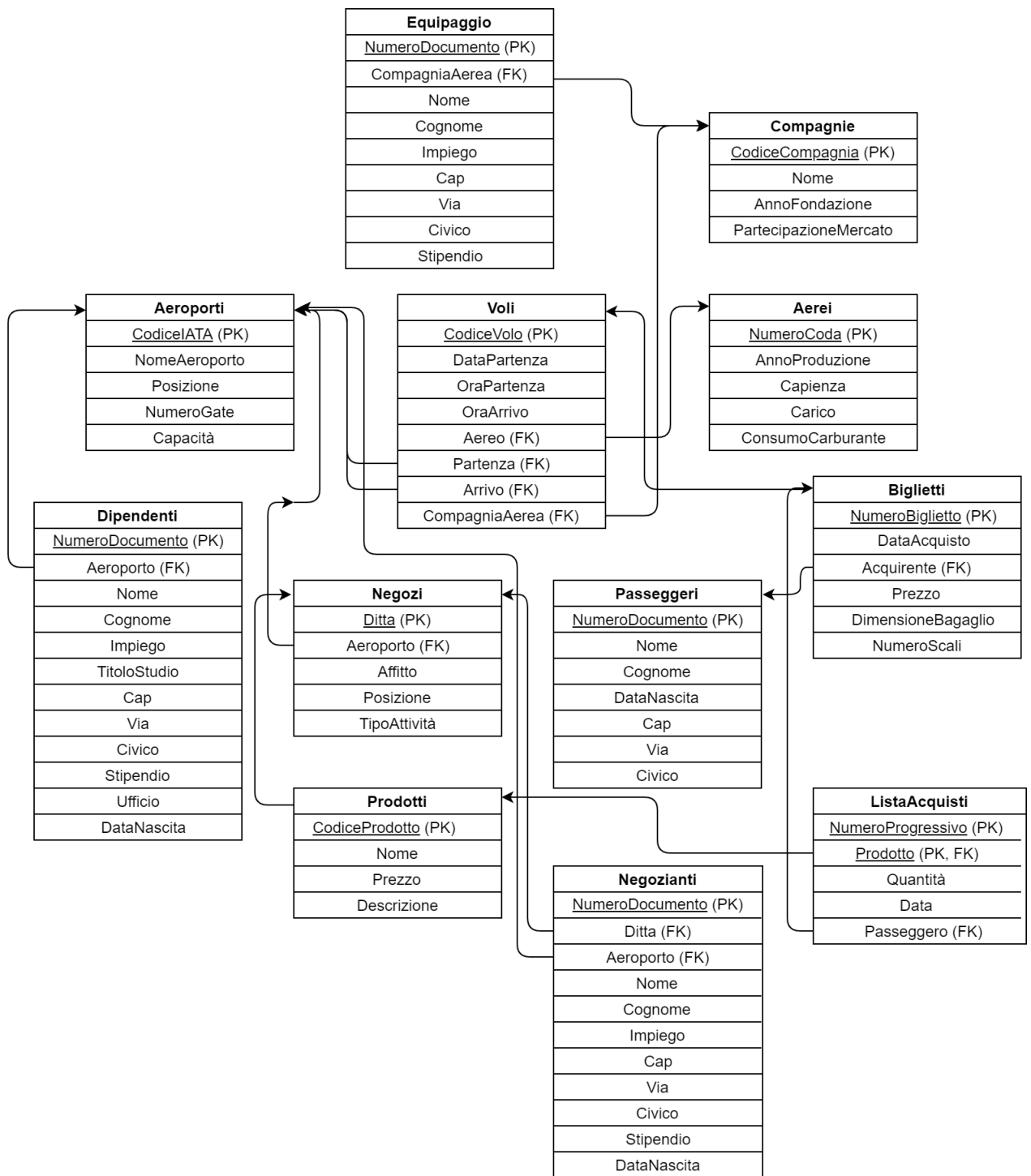
### 3 Dizionario dei dati e primo schema logico

Entità	Descrizione	Attributi	Identificatore
Aeroporto	Aeroporto italiano ENAC (che ospita voli di linea)	Nome, regione geografica, numero di gate e capacità	Codice IATA
Volo	Tratta aerea tra due aeroporti	Data di partenza, data di arrivo, ora di partenza, ora di arrivo, aereo che ha eseguito il volo, aeroporto di partenza, aeroporto di arrivo, compagnia aerea	Codice di volo
Aereo	Mezzo di trasporto che esegue il volo	Anno di produzione, capienza, carico, consumo di carburante	Numero di coda
Compagnia aerea	Impresa adibita al trasporto di persone mediante l'utilizzo di aerei	Nome, anno di fondazione e quota di partecipazione di mercato	Codice IATA (per compagnie aeree)
Equipaggio	Personale che lavora sull'aereo	Compagnia aerea, nome, cognome, impiego, indirizzo, stipendio	Numero di documento
Biglietto	Biglietto necessario per prendere l'aereo	Voli, acquirente, prezzo, dimensione del bagaglio prevista, numero di scali, data di acquisto	Numero di biglietto
Passeggero	Persona che compra un biglietto aereo	Nome, cognome, data di nascita, indirizzo	Numero di documento
Negozi	Attività commerciale nel duty-free	Aeroporto, inizio e durata del contratto d'affitto, posizione, tipo di attività	Ditta
Prodotto	Oggetto di vendita nei negozi	Nome, ditta, descrizione e prezzo	Codice del prodotto
Lista acquisti	Lista degli acquisti effettuati dai passeggeri in aeroporto	Numero di scontrino, prodotto acquistato, quantità, data di acquisto e passeggero	Numero progressivo e prodotto
Dipendente	Lavoratore dipendente presso l'aeroporto	Nome, cognome, indirizzo, data di nascita, impiego, titolo di studio, stipendio, ufficio	Numero di documento

Dizionario delle entità

Relazione	Descrizione	Composizione	Cardinalità
Lavoro	Relazione che lega un lavoratore ad un luogo di lavoro	Dipendenti - Aeroporti, Negozianti - Negozi, Equipaggio - Compagnie	(1,1) - (0,N)
Proprietà	Relazione che lega ciascun volo alla compagnia aerea che lo gestisce	Voli - Compagnie	(1,1) - (0,N)
Mezzo	Relazione che lega un volo all'aereo che compierà il volo	Voli - Aerei	(1,1) - (0,N)
Partenza	Relazione che lega un volo all'aeroporto di partenza	Voli - Aeroporti	(1,1) - (0,N)
Arrivo	Relazione che lega un volo all'aeroporto di arrivo	Voli - Aeroporti	(1,1) - (0,N)
DutyFree	Relazione che colloca ciascun negozio nel duty-free di un aeroporto	Negozi - Aeroporti	(1,1) - (0,N)
Vendita	Lega un prodotto ai negozi in cui viene messo in vendita	Prodotti - Negozi	(1,N) - (0,N)
Acquisto	Associa alla lista acquisti il prodotto acquistato	ListaAcquisti - Prodotti	(1,1) - (0,N)
Tracciamento	Associa ad un acquisto della lista acquisti il biglietto	ListaAcquisti - Biglietti	(1,1) - (0,N)
Intestatario	Associa a ciascun biglietto il passeggero che l'ha acquistato	Biglietti - Passeggeri	(1,1) - (1,N)
Imbarco	Associa a ciascun biglietto i voli previsti	Biglietti - Voli	(1,N) - (0,N)

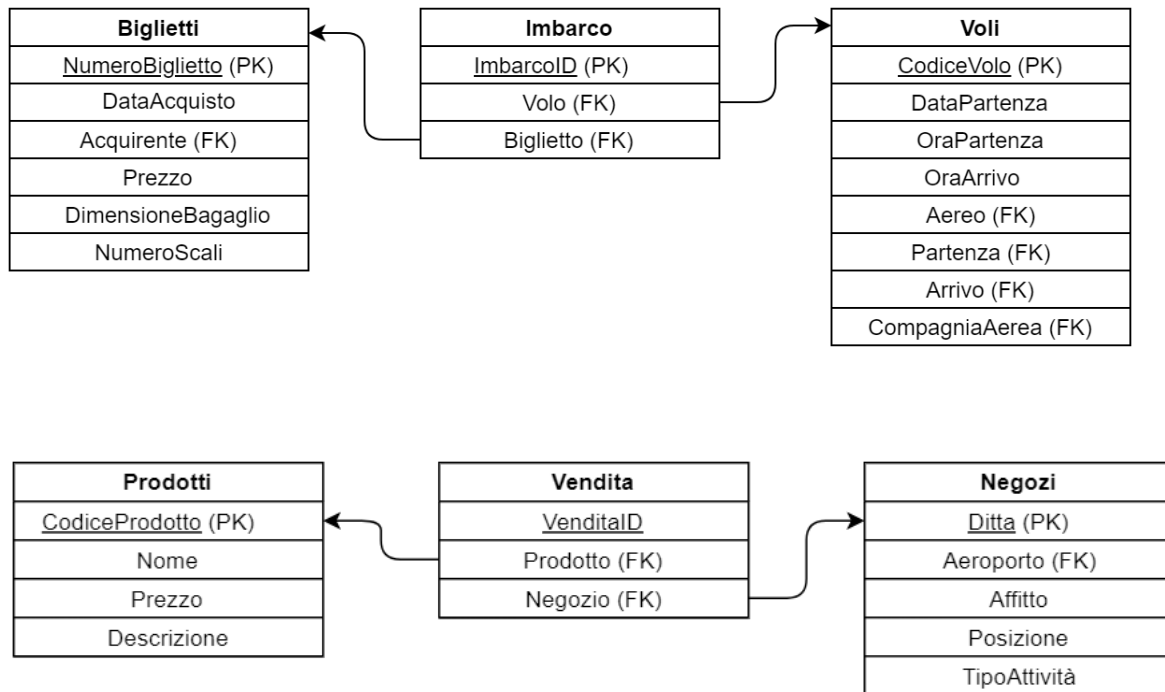
Dizionario delle relazioni



Schema logico

## Criticità

Nello schema logico emerge immediatamente una criticità; sono presenti tutte le tabelle cui siamo interessati, tuttavia ci sono due collegamenti tra tabelle che non sono chiari: si tratta del collegamento tra la tabella Prodotto e la tabella Negozi (relazione chiamata "Vendita") e del collegamento tra la tabella Biglietti e la tabella Voli (relazione chiamata "Imbarco"). Nello schema è quindi necessario introdurre due nuove tabelle che serviranno unicamente a gestire queste due relazioni.



Correzioni schema logico

## Identificatori

- **Aeroporti**: L'identificatore IATA è un codice univoco di tre lettere che identifica ogni aeroporto del mondo.
- **Voli**: Il codice di volo è una combinazione di una parte alfabetica indicante la compagnia aerea e una parte numerica indicante il numero di volo.
- **Aerei**: Il numero di coda è una sigla alfanumerica che identifica gli aeromobili come la targa per le autovetture.
- **Compagnie**: Il codice della compagnia è una combinazione di lettere univoca per ciascuna compagnia aerea.
- **Equipaggio**: Il numero di documento di identità.
- **Biglietti**: Il numero di biglietto è un numero di 13 cifre identificativo di ciascun biglietto.
- **Passeggeri**: Il numero di documento di identità.
- **Lista Acquisti**: Il numero progressivo è il numero dello scontrino, identificativo congiuntamente al prodotto (attraverso il codice del prodotto) indicato sullo scontrino.
- **Prodotti**: Il codice prodotto EAN è un codice numerico identificativo di ogni prodotto soggetto alla vendita al dettaglio in Unione Europea.

- **Negozi:** Il codice ditta è univoco su scala nazionale (in caso di estensione del database va rivisto).
- **Dipendenti:** Il numero di documento di identità.
- **Negozianti:** Il numero di documento di identità.

## 4 Analisi delle ridondanze

Nel database sono presenti delle ridondanze che possono essere eliminate:

- La prima ridondanza non si verifica necessariamente, ma solo se ogni biglietto è venduto esclusivamente sul territorio nazionale. Si tratta del numero di scali, attributo del biglietto aereo: il numero di scali, infatti, è il numero di volte che un passeggero deve atterrare in un aeroporto diverso dalla destinazione prima di arrivare. Pertanto, il numero di scali si può tradurre come il numero di voli previsti da un biglietto aereo  $-1$ . Si osserva infatti come è possibile conoscere il numero di scali del biglietto numero 4449993331111 (numero di biglietto fittizio di 13 cifre) nei due seguenti modi, che danno lo stesso output:

```
SELECT COUNT(Volo)-1
FROM Imbarco
WHERE Biglietto = 4449993331111;
```

```
SELECT NumeroScali
FROM Biglietti
WHERE NumeroBiglietto = 4449993331111;
```

- La seconda ridondanza è meno evidente ed è identificabile nell'attributo "CompagniaAerea" nella tabella "Voli": infatti, è possibile evincere la compagnia aerea che gestisce un volo dal suo codice di volo, in quanto le prime due lettere del codice identificativo di un volo indicano la compagnia aerea che lo gestisce. Anche in questo caso, segue un esempio in cui visualizziamo la compagnia aerea del volo "FR5225" in due modi diversi:

```
SELECT c.Nome AS NomeCompagnia, v.CodiceVolo
FROM Compagnie c
LEFT JOIN Voli v
ON CompagniaAerea = CodiceCompagnia
WHERE CodiceVolo = 'FR5225';
```

```
SELECT c.Nome AS NomeCompagnia, v.CodiceVolo
FROM Compagnie c
LEFT JOIN Voli v
ON LEFT(CodiceVolo,2) = CodiceCompagnia
WHERE CodiceVolo = 'FR5225';
```

Pertanto è possibile concludere che le variabili "NumeroScali" nella tabella "Biglietti" e "NomeCompagnia" nella tabella "Voli" possono essere eliminate. Un discorso analogo non si può fare per la variabile "NomeCompagnia" nella tabella "Equipaggio", in quanto è necessario conoscere il codice della compagnia in modo esplicito non avendo altre informazioni che ci permettano di estrapolarlo come il codice del volo.

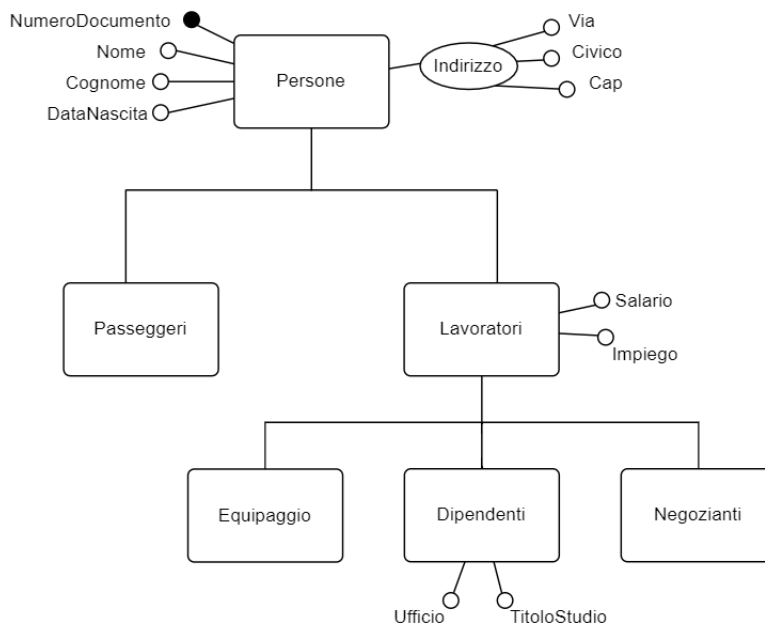
## 5 Generalizzazioni

Nel database sono presenti più tabelle che fanno riferimento a delle persone e richiedono gli stessi dati anagrafici, pertanto è possibile stabilire una relazione di ereditarietà tra le entità. In particolare, si può dire che esiste un'entità genitore "Persone" e che le sue proprietà sono ereditate dalle entità "Passeggeri", "Negozianti", "Dipendenti" e "Equipaggio".

Analogamente è possibile individuare un'ulteriore entità genitore, ovvero "Lavoratori": infatti questa entità potrebbe ereditare le proprietà di "Persone" e indicare ulteriormente anche l'impiego e lo stipendio, facendo ereditare questi due attributi a "Negozianti", "Dipendenti" e "Equipaggio".

La generalizzazione "Lavoratori" può considerarsi esclusiva, assumendo che non ci siano persone che svolgono due lavori entrambi nell'ambito dei voli aerei: infatti se un lavoratore appartiene ad una delle entità figlie, non apparterrà alle altre; Al contrario, la generalizzazione "Persone" non può considerarsi tale: un lavoratore, indipendentemente da quale categoria appartenga, può acquistare un biglietto aereo e viaggiare, pertanto "Persone" risulta essere una generalizzazione sovrapposta.

Entrambe le generalizzazioni, inoltre, possono considerarsi parziali, poiché ci possono essere persone che non hanno né acquistato un biglietto, né lavorano; allo stesso modo possono esserci lavoratori che non fanno parte delle tre categorie di lavoro sottostanti. Tuttavia è possibile fare una considerazione: essendo questo database mirato a contenere dati utili per informazioni di volo, non ha senso pensare l'entità "Persone" come un'entità che contenga come istanze tutte le persone in generale, ma soltanto quelle interessate a noi, ovvero quelle che appartengono o alla categoria passeggero o lavoratori; può, di conseguenza, essere considerata una generalizzazione totale. Un discorso analogo si può fare anche per l'entità "Lavoratori".



Ereditarietà



## 6 Vincoli non esprimibili

- La data di arrivo di un volo deve essere la stessa o seguente alla data di partenza e, in caso siano uguali, l'ora di arrivo deve essere seguente all'ora di partenza. Può essere gestito con un trigger:

```
DELIMITER $$
CREATE TRIGGER check_date_voli
BEFORE INSERT ON Voli
FOR EACH ROW
BEGIN
    IF NEW.DataArrivo < NEW.DataPartenza THEN
        SIGNAL SQLSTATE '45001' SET MESSAGE_TEXT = 'La data di arrivo deve essere
            successiva o uguale alla data di partenza';
    ELSEIF NEW.DataArrivo = NEW.DataPartenza
        AND NEW.OraArrivo <= NEW.OraPartenza THEN
        SIGNAL SQLSTATE '45001' SET MESSAGE_TEXT = "L\'ora di arrivo deve essere
            % successiva all\'ora di partenza se la data è la stessa";
    END IF;
END $$
DELIMITER ;
```

- Se un biglietto ti permette di effettuare  $n$  voli, qualsiasi biglietto che ti permette di fare un numero di voli  $< n$  compresi negli  $n$  voli del primo biglietto, deve avere un prezzo minore (anche se per ragioni di mercato di fatto non è sempre così).  
es. Con il biglietto  $A$  posso fare Trieste - Parigi facendo scalo a Roma, allora il biglietto  $B$  che mi fa fare Trieste - Roma sullo stesso volo del biglietto  $A$  avrà un prezzo minore di  $A$ .

## 7 Tabella dei volumi

Si considerano i 41 aeroporti italiani di linea ENAC, i quali mediamente ospitano in un giorno 350 voli nazionali. Considero un periodo di tempo di un anno.

Concetto	Tipo	Volume
Aeroporti	E	41
Voli	E	127750
Passeggeri	E	9.000.000
Aerei	E	8.000
Compagnia aerea	E	15
Equipaggio	E	10.000
Biglietti	E	10.000.000
Lista Acquisti	E	15.000.000
Negozi	E	800
Negozianti	E	7.000
Prodotti	E	80.000
Dipendenti	E	25.000
Lavoro	R	$10.000 + 25.000 + 7.000$
Proprietà	R	127.750
Mezzo	R	127.750
Partenza	R	127.750
Arrivo	R	127.750
DutyFree	R	800
Imbarco	R	12.000.000
Intestatario	R	10.000.000
Tracciamento	R	15.000.000
Acquisto	R	15.000.000
Vendita	R	90.000

Tabella dei Volumi

## 8 Normalizzazione

- La prima forma normale è rispettata, in quanto ogni attributo assume solamente valori atomici.
- La seconda forma normale è rispettata, poiché tutti gli attributi nelle varie tabelle dipendono funzionalmente dalla chiave primaria.
- La terza forma normale è rispettata, poiché tutti gli attributi dipendono esclusivamente dalla chiave primaria e non dagli altri attributi non-chiave.

## 9 Operazioni

1. Visualizzare quanti voli sono stati gestiti da ciascuna compagnia durante l'anno:

```
DELIMITER $$
CREATE PROCEDURE VoliPerCompagnia()
BEGIN
    SELECT c.Nome AS Compagnia, COUNT(v.CodiceVolo) AS NumeroVoli
    FROM Compagnie c
    LEFT JOIN Voli v ON c.CodiceCompagnia = v.CompagniaAerea
    GROUP BY c.Nome
    ORDER BY NumeroVoli DESC;
END $$
DELIMITER ;
```

2. Visualizzare chi ha viaggiato mediamente più di una volta al mese (ricordando che il database fa riferimento ad un anno):

```
CREATE PROCEDURE GrandiViaggiatori()
BEGIN
    SELECT p.Nome, p.Cognome, COUNT(*) AS NumeroViaggi
    FROM Persone p
    INNER JOIN Passeggeri pas ON p.NumeroDocumento = pas.NumeroDocumento
    INNER JOIN Biglietti b ON pas.NumeroDocumento = b.Acquirente
    GROUP BY p.Nome, p.Cognome
    HAVING COUNT(*) > 12;
END $$
DELIMITER ;
```

3. Visualizzare le procedure di chi ha effettivamente viaggiato almeno una volta ogni mese (non in media):

```
DELIMITER $$
CREATE PROCEDURE ViaggiatoriPeriodici()
BEGIN
    SELECT p.Nome, p.Cognome
    FROM Persone p
    INNER JOIN Passeggeri pas ON p.NumeroDocumento = pas.NumeroDocumento
    INNER JOIN Biglietti b ON b.Acquirente = pas.NumeroDocumento
    INNER JOIN Imbarco i ON b.NumeroBiglietto = i.Biglietto
    INNER JOIN Voli v ON i.Volo = v.CodiceVolo
    GROUP BY p.Nome, p.Cognome
    HAVING COUNT(DISTINCT DATE_FORMAT(v.DataPartenza, '%Y-%m')) >=12;
END $$
DELIMITER ;
```

4. Visualizzare il nome e il titolo di studio dei dipendenti degli aeroporti che prendono più di una cifra predefinita al mese:

```
DELIMITER $$
CREATE PROCEDURE StipendioAlto(IN soldi INT)
BEGIN
    SELECT p.Nome, d.TitoloStudio
    FROM Persone p
    INNER JOIN Dipendenti d ON p.NumeroDocumento = d.NumeroDocumento
    WHERE d.Stipendio >= soldi;
END $$
DELIMITER ;
```

5. Visualizzare il prodotto più venduto:

```
DELIMITER $$
CREATE PROCEDURE ProdottoPiùVenduto()
BEGIN
    SELECT p.Nome, p.Prezzo, SUM(l.Quantita) AS NumeroVendite
    FROM Prodotti p
    INNER JOIN ListaAcquisti l ON p.CodiceProdotto = l.Prodotto
    GROUP BY p.Nome, p.Prezzo
    ORDER BY NumeroVendite DESC
    LIMIT 1;
END $$
DELIMITER ;
```

6. Visualizzare le 5 regioni con più traffico aereo:

```
DELIMITER $$
CREATE PROCEDURE RegioniTrafficate()
BEGIN
    SELECT a.Posizione, COUNT(v1.Partenza) + COUNT(v2.Arrivo) AS NumeroVoli
    FROM Aeroporti a
    INNER JOIN Voli v1 ON a.CodiceIata = v1.Partenza
    INNER JOIN Voli v2 ON a.CodiceIata = v2.Arrivo
    GROUP BY a.Posizione
    ORDER BY NumeroVoli DESC
    LIMIT 5;
END $$
DELIMITER ;
```