What US airport has seen the highest growth in passengers in the 1990 to 2009 period?

4th December 2021

Origin Data Set Description

■ The origin data set is over 3.5 million monthly domestic flight records from 1990 to 2009. Data are arranged as an adjacency list with metadata.

Short name	Type	Description
Origin	String	Three letter airport code of the origin airport
Destination	String	Three letter airport code of the destination airport
Origin City	String	Origin city name
Destination City	String	Destination city name
Passengers	Integer	Number of passengers transported from origin to destination
Seats	Integer	Number of seats available on flights from origin to destination
Flights	Integer	Number of flights between origin and destination (multiple records for one month, many with flights > 1)
Distance	Integer	Distance (to nearest mile) flown between origin and destination
Fly Date	Integer	The date (yyyymm) of flight
Origin Population	Integer	Origin city's population as reported by US Census
Destination Population	Integer	Destination city's population as reported by US Census

##Snippet:

```
MFR    RDM    Medford, OR       Bend, OR       0     0     1    156    200810    200298    157730
AMA    EKO    Amarillo, TX     Elko, NV      124   124    1    858    199308    202960     40259
TUS    EKO    Tucson, AZ       Elko, NV      112   124    1    658    199308    711392     40259
AMA    EKO    Amarillo, TX     Elko, NV      115   124    1    858    199406    206315     41668
ICT    EKO    Wichita, KS      Elko, NV      100   124    1   1007    199607    552884     45034
SPS    EKO    Wichita Falls, TX   Elko, NV   122   124    1   1059    199603    147683     45034
```

Data Preparation

```python
In [132]: flights = pd.read_csv ('flight_edges.tsv', sep='\t', names=["Origin", "Destination", "Origin City", "Destination Cit
```

```python
In [133]: flights.head()
```

Out[133]:

	Origin	Destination	Origin City	Destination City	Passengers	Seats	Flights	Distance	Fly Date	Origin Population	Destination Population
0	MHK	AMW	Manhattan, KS	Ames, IA	21	30	1	254.0	200810	122049	86219
1	EUG	RDM	Eugene, OR	Bend, OR	41	396	22	103.0	199011	284093	76034
2	EUG	RDM	Eugene, OR	Bend, OR	88	342	19	103.0	199012	284093	76034
3	EUG	RDM	Eugene, OR	Bend, OR	11	72	4	103.0	199010	284093	76034
4	MFR	RDM	Medford, OR	Bend, OR	0	18	1	156.0	199002	147300	76034

```python
In [134]: flights.shape
```

Out[134]: (3606803, 11)

```python
In [135]: flights.describe(include='all').loc['unique', :]
```

```
Out[135]: Origin                    683
          Destination               708
          Origin City               535
          Destination City          548
          Passengers                NaN
          Seats                     NaN
          Flights                   NaN
          Distance                  NaN
          Fly Date                  NaN
          Origin Population         NaN
          Destination Population    NaN
          Name: unique, dtype: object
```

```python
In [139]: flights_small = flights.sample(1000)
```

```python
In [140]: flights_small.head()
```

Out[140]:

	Origin	Destination	Origin City	Destination City	Passengers	Seats	Flights	Distance	Fly Date	Origin Population	Destination Population	Fly Year	Fly Month
2568185	IAD	GSO	Washington, DC	Greensboro, NC	975	2953	77	239.0	199111	8369110	548939	1991	11
1712121	DFW	LGA	Dallas, TX	New York, NY	41891	64098	462	1389.0	199807	9475022	34683010	1998	07
670670	TUS	ATL	Tucson, AZ	Atlanta, GA	8157	9514	67	1541.0	200602	975476	5119641	2006	02
3103072	DTW	MSP	Detroit, MI	Minneapolis, MN	6590	7400	74	528.0	200307	8985512	3078253	2003	07
1877232	LAX	SYR	Los Angeles, CA	Syracuse, NY	123	168	1	2351.0	199707	23583460	656224	1997	07

First results

	Passengers
Fly Year	
1990	370404412
1991	366689941
1992	383323302
1993	391096541
1994	426130721
1995	441706772
1996	467885967
1997	481929580
1998	491917421
1999	511338939
2000	533957069
2001	497009692
2002	486028550
2003	508988965
2004	546469416
2005	568207371
2006	564727445
2007	579257485
2008	554998719
2009	526301909

Total number of passengers, yearly

```
In [197]: flights.pivot_table(index=["Fly Year"], values=["Passengers"], aggfunc=np.sum).plot(kind="bar")

Out[197]: <AxesSubplot:xlabel='Fly Year'>
```